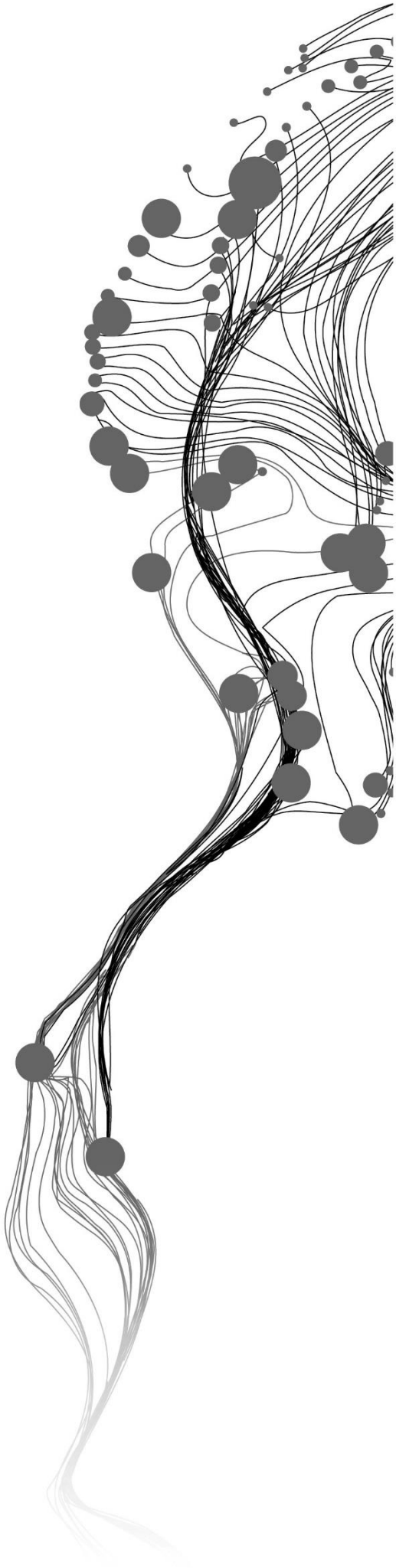
EXTRACTING CADASTRAL BOUNDARIES FROM UAV IMAGES USING FULLY CONVOLUTIONAL NETWORKS

XUE XIA
February, 2019

SUPERVISORS:
Dr. M.N. Koeva
Dr. C. Persello



EXTRACTING CADASTRAL BOUNDARIES FROM UAV IMAGES USING FULLY CONVOLUTIONAL NETWORKS

XUE XIA
Enschede, The Netherlands, February, 2019

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Urban Planning and Management

SUPERVISORS:
Dr. M.N. Koeva
Dr. C. Persello

THESIS ASSESSMENT BOARD:
Prof.dr. C.H.J. Lemmen (Chair)
Dr. R.M. Bennett (External Examiner, University of Reading)

ABSTRACT

It is estimated that 70% of the land rights in the world remain unregistered. Traditional field surveying approach is costly and labour intensive. Imagery-based cadastral mapping along with automatic feature detection techniques are recently being exploited by researchers to accelerate this process. Most prior studies are explored on rural or agricultural areas with relatively simple morphology and clear boundaries, based on image segmentation or edge detection. Our research seeks to propose a novel strategy for cadastral boundary detection in complexed urban environments. We introduce a technique based on deep Fully Convolutional Networks (FCNs), which can automatically learn high-level spatial features from images, to extract cadastral boundaries. Our strategy combines FCNs and a grouping algorithm using the Oriented Watershed Transform (OWT) to generate connected contours. These contours are presented hierarchically in an Ultrametric Contour Map (UCM). We call this workflow of producing connected cadastral boundaries FCN-OWT-UCM.

FCN-OWT-UCM is tested in two case study locations in Rwanda, Busogo and Muhoza, using images acquired by Unmanned Aerial Vehicles (UAV) in 2018. The performance in Muhoza is slightly better than Busogo. An average of 0.47 in precision (correctness), 0.52 in recall (completeness) and 0.49 in F-score (the harmonic mean between precision and recall) are achieved. The proposed method is compared with state-of-the-art Multi-Resolution Segmentation (MRS) and Globalized Probability of Boundary (gPb) in both study sites and gains better overall accuracy. In conclusion, FCN-OWT-UCM is able to effectively extract urban cadastral boundaries, the accuracy of which largely depends on the proportion of visible boundaries. This automated method could minimize manual digitization and reduce field work, thus facilitating the current cadastral mapping and updating practices.

Keywords: deep learning, fully convolutional networks, cadastral boundaries, contour detection, unmanned aerial vehicles.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my super supervisors Dr. Mila Koeva and Dr. Claudio Persello. Thank you so much for your wholehearted support and constructive guidance. You are always a team, erudite, motivated and hopeful. Important things are not easy, so is the MSc. thesis. Your spirit encouraged me a lot, teaching me to stay cool, be strong and brave. My sincere gratitude also goes to Prof. Chrit Lemmen. Your valuable suggestions during the proposal and midterm phase contributed to shape my research. I also would like to thank Dr. Divyani Kohli. I appreciate your generous help in eCognition.

Special thanks go to ITC Excellence programme for providing me the scholarship to study at ITC. It is a fantastic experience. I am deeply grateful to the entire ITC family. So many warm and cute faces, love you all!

Finally, I would like to thank my family. My dear mother, father and little brother, your love keeps me going. Thank you and I love you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background and justification

Cadastres, which record the physical location and ownership of the real properties, are the basis of land administration systems (Luo, Bennett, Koeva, & Lemmen, 2017). Nowadays, cadastral mapping has received considerable critical attention. An effective cadastral system formalises private property rights, which is very important to promote agricultural productivity, secure effective land market, reduce poverty and support national development in the broadest sense (Williamson, 1997). However, it is estimated that over 70% people in this world do not have access to a formal cadastral system (Enemark et al., 2016). The key obstruction is the use of traditional field surveying approach to record land parcels, which is often claimed to be time-consuming, costly and labour intensive.

In response to the problem, Enemark, Bell, Lemmen, and Mclaren (2014) proposed fit-for-purpose (FFP) land administration, suggesting that land administration systems should be designed to meet the current needs of a specific country, rather than following strict technical standards. One of the key principles of the FFP approach is using 'general' rather than 'fixed' boundaries. Boundaries measured through on-site cadastral survey using total station or Global Navigation Satellite System (GNSS) with a precise location are considered as fixed, while boundaries delineated from high-resolution imagery are called general (Crommelinck et al., 2016). General boundary approaches extract cadastral boundaries by visually interpreting and manually digitizing from georeferenced high-resolution imageries. In this regard, huge fieldwork surveying tasks could be avoided. Therefore, although less spatially precise, general boundary approaches are much cheaper and faster than conventional cadastral survey. Typically, high-resolution satellite images (HRSI) have been used for interpreting general boundaries, but there are still obstacles such as high-cost, cloudy or dated imagery (Ramadhani, Bennett, & Nex, 2018).

Nowadays, Unmanned Aerial Vehicles (UAV) are increasingly being used as a low-cost, affordable platform which can support in acquiring high-resolution data. UAVs can fly under clouds and capture sub-meter level imagery in a cheap and timely fashion. Several studies already focus on using UAVs for cadastral mapping. Cunningham, Walker, Stahlke, Wilson and Opportunity (2011) used UAVs as a substitute of field surveys for cadastral mapping in rural Alaska. From their experiences, property lines could be well defined through meetings with involved stakeholders using the accurate orthophoto. Rubinov, Biraro, Fuller and Bennett (2015) claimed that UAVs have the potential to revolutionise land administration activities for its flexible maneuverings and supplying of high-quality, high-resolution imagery. Koeva, Muneza, Gevaert, Gerke, and Nex (2018) introduced the whole workflow of using UAVs for map creation and updating, consisting of flight planning, data acquisition, image processing and orthophoto creation.

To further expediting the mapping process, automatic boundary detection techniques from UAV images are being explored by researchers. In practice, cadastral boundaries are often marked by physical objects, such as roads, building walls, fences, water drainages, ditches, rivers etc. (Luo et al., 2017). Such boundaries are visible in UAV images and bear the potential to be automatically extracted through image analysis algorithms. However, due to the inherent complexity of cadastral boundary, automatic detection often shows a certain error rate and needs to be manually corrected through post-processing steps

(Crommelinck et al., 2016). Despite the needs for post-editing, these semi-automatic approaches still turn out to be cheaper and faster than digitizing from scratch for all boundaries.

The most commonly used techniques for semi-automatic boundary delineation are based on image segmentation and edge detection (Crommelinck et al., 2016). Segmentation refers to partitioning images into disjoint regions, inside which the pixels are similar to each other with regard to spectral characteristics (Pal & Pal, 1993). On the other hand, edge detection techniques model edges as sharp discontinuities in brightness and colour (Bowyer, Kranenburg, & Dougherty, 2001). These techniques are based on low-level features and are not accurate enough. For cadastral boundary detection, the semantic gap between high-level boundary concept and low-level visual cues forms the main challenge. More reliable and informative features should be constructed to bridge the semantic gap, thus more advanced feature extraction techniques are needed.

Recent studies indicate that deep learning methods such as Convolutional Neural Networks (CNNs) are highly effective for extraction of higher-level representations needed for detection or classification from raw input (Zhu et al., 2017), which brings in new opportunities in cadastral boundary detection. Fully Convolutional Networks (FCNs) are a more recent deep learning method. In contrast to traditional CNNs, which predict only the label of the central pixel of the image, FCNs directly perform pixel-wise predictions. The input image of an FCN can have arbitrary size, and all the pixels of the input will be labelled. Therefore, compared to CNNs, FCNs can largely reduce computational cost and processing time. The superiority of FCNs in feature learning and computational efficiency makes them promising for the detection of visible cadastral boundaries. However, very little research uses FCNs in extraction of cadastral boundary. To the best of the author's knowledge, only Musyoka (2018) applied FCNs in agricultural field boundaries delineation and achieved better accuracy and visual quality compared to Multi-resolution segmentation, globalized probability of boundary (gPb) and Canny detector.

What's more, most automatization for boundary detection was explored on rural areas or agricultural fields with very clearly visible boundaries, including the prominent work by Babawuro and Beiji (2012), Nyandwi (2018), Wassie, Koeva, Bennett and Lemmen (2018), Musyoka (2018) etc.. Urban areas with more complexed morphology are very challenging for standard techniques such as image segmentation and edge detection. The former is in high risk of over-segmentation as there are often many different objects within one land parcel. For the latter, all the outlines of these objects will be detected by edge detectors, leading to many false detections.

Taking above challenges and opportunities into consideration, this study aims at using fully convolutional networks for extracting cadastral boundaries from UAV images. Special focus is given to urban and sub-urban areas. The detection results of FCNs are fragmented boundaries. A grouping method to connect the disjoint boundaries automatically is also investigated in this research by using Oriented Watershed Transform (OWT) (Arbeláez, Maire, Fowlkes, & Malik, 2011). The final result is presented as an Ultrametric Contour Map (UCM), representing the connected boundaries at different levels of detail.

1.2. Research problem

There is a clear need for innovative and automatic cadastral boundary delineation methods to face the challenge of 70% unregistered land rights. Much effort is already put into automatic techniques such as image segmentation and edge detection. However, these methods are mostly applied in rural areas with simple landforms and clear boundaries. Furthermore, there is still need to improve the accuracy and computational efficiency. Recent research shows that deep learning has great ability in learning features from raw image, but very limited research uses FCNs in cadastral boundary detection, especially in urban

region. Therefore, extracting cadastral boundaries in urban and sub-urban environment using FCNs from UAV images forms the problem of this research. To connect the disjoint boundaries, a workflow of FCN-OWT-UCM is introduced to further modify the result.

1.3. Research objectives and questions

1.3.1. General objectives

To extract cadastral boundaries from UAV images using FCNs.

1.3.2. Specific objectives

Objective 1: To prepare the data for cadastral boundary detection.
Objective 2: To develop a methodology for cadastral boundary detection.
Objective 3: To compare the performance of the proposed method with other state-of-the-art methods.

1.3.3. Research questions

Objective 1:
1. How to prepare the reference data?
2. How to design the training and testing dataset?

Objective 2:
1. Which FCN architecture is appropriate for boundary detection?
2. What are the optimal hyper-parameters for the proposed FCN?
3. How to improve the result of FCN?

Objective 3:
1. What are the widely accepted metrics for evaluating boundary detection?
2. Which method achieves the best result?

1.4. Conceptual framework

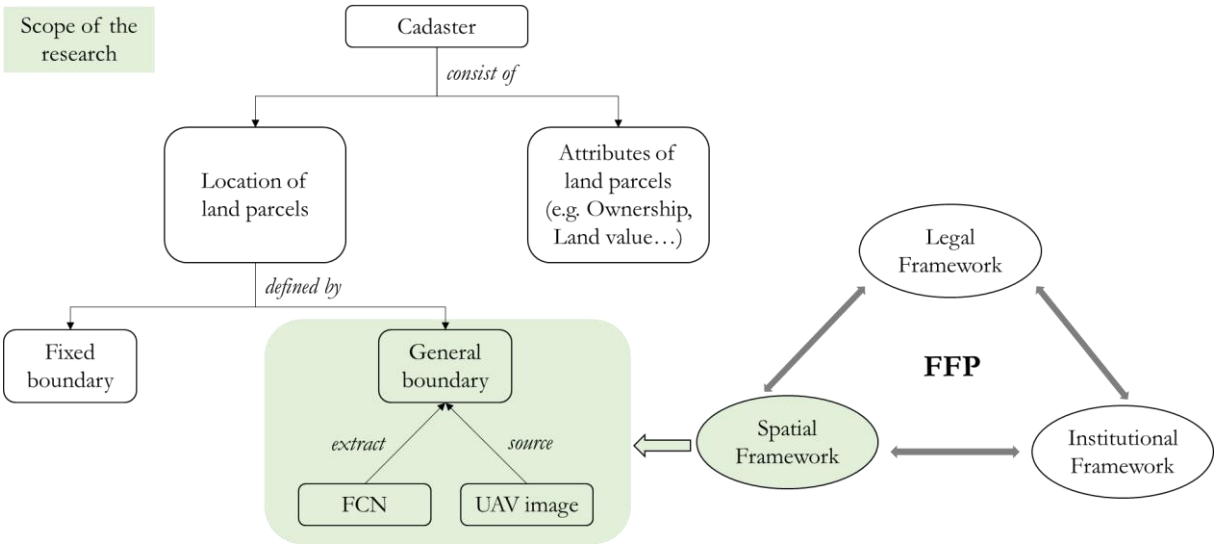


Figure 1: Conceptual framework. The concepts marked with green form the scope of this research.

The conceptual framework is presented in Figure 1. Frames and texts marked by green colour define the scope of this research. The context-specific aim is general cadastral boundary mapping. The term cadastral boundary is rooted in cadastre, which records the location and attributes of land parcels. The latter often include information on ownership, land value, etc.. The location of land parcels can be defined by fixed

boundary or general boundary. In line with FFP land administration, the potential for general boundary detection is explored in this research. The concept of FFP land administration contains three core frameworks, including spatial, legal and institutional aspects. These three components correspond to geographical, judicial and administrative context respectively and they are interrelated (Enemark et al., 2016). Our research is under the context of spatial framework by involving UAV images and FCNs in cadastral mapping.

1.5. Thesis structure

The structure of this thesis is organized as following:

Chapter 1. Introduction

This chapter gives the background and justification of the research, clarifying the research problem, objectives and questions. The main concepts and the underlying relations are indicated in a conceptual framework.

Chapter 2. Literature review

Related concepts for cadastral boundary and state-of-the-art feature extraction techniques are reviewed in this chapter. Former relevant scientific literature is also reviewed in this part.

Chapter 3. Methodology

An overview of the research methodology and the study areas are introduced in this chapter, followed by a detailed description of each step, including data preparation, automation process for boundary detection, alternative approaches and accuracy assessment.

Chapter 4. Results and analysis

The experimental results are presented here with a brief explanation. The results are presented in the sequence of hyper-parameter tuning result for FCN, threshold tuning result for OWT-UCM and the accuracy assessment result for final implementation. The results of alternative approaches are also described and compared in this chapter.

Chapter 5. Discussion

In this chapter, a elaborate discussion for the obtained results is presented.

Chapter 6. Conclusions and recommendations

This chapter closes the thesis with concluding remarks of the whole research and recommendations for future study.

1.6. Summary

This chapter introduces the background of this research in social and academical aspects, leading to the actual research problem. The scope of this research and the overall structure of the thesis are illustrated to provide a global perception. In summary, this study aims to extract general cadastral boundaries following FFP land administration principles. The innovation of this research is in using FCNs and the application of cadastral boundary extraction in urban environments.

2. LITERATURE REVIEW

2.1. Concepts related to cadastral boundaries

The term cadastre has various definitions. United Nations Economic Commission for Europe (UNECE) defined cadastre as an information system which consists of two parts: "a series of maps or plans showing the size and location of all land parcels together with text records that describe the attributes of the land" (UNECE, 1996, p.11). Silva and Stubkjær (2002, p.410) described cadastre as "a systematic and official description of land parcels, which includes for each parcel a unique identifier". Luo et al. (2017, p.2) suggested that cadastre is "a comprehensive official record of the real property's boundary, ownership, and value, and is a systematically arranged public register that depends on survey measurements". According to these definitions, land parcels are an essential geospatial component of a cadastre.

The land parcels are defined by cadastral boundaries. There are direct and indirect cadastral surveying techniques for the acquisition of cadastral boundaries. Direct techniques measure the accurate position of a boundary through on the ground measurement; indirect techniques are those relying on remotely sensed data instead of going to the field (Crommelinck et al., 2016). Depending on whether the boundary is accurately surveyed and determined, cadastral boundaries can be defined as fixed or general (Luo et al., 2017). Generally speaking, fixed boundaries are commonly measured by direct techniques, or some indirect techniques using very high-resolution data that meet the required accuracy; general boundaries are more likely to be visible boundaries extracted from remotely sensed data (Crommelinck et al., 2016). Rather than acquiring fix boundaries following advanced technical standards at a high cost, fit-for-purpose land administration seeks to support more rapid and low-cost cadastral mapping to safeguard people's land rights. The underlying truth is that provision of tenure security does not necessarily require centimetre-accurate cadastral boundaries (Wassie et al., 2018). Therefore, this study focuses on delineating general boundaries using indirect surveying techniques.

2.2. Prior studies for imegery-based boundary detection

Imagery-based cadastral mapping is in recent years being explored by researchers for the purpose of quick and cost-effective data acquisition and updating. Manual digitization along with stakeholder consultation was conducted for imagery-based boundary detection in earliest practises (Manyoky, Theiler, Steudler, & Eisenbeiss, 2012; Ali & Ahmed, 2013; Parida, Sanabada, & Tripathi, 2014). According to these case studies, it can be concluded that more land parcels can be surveyed in less time by using high-resolution imagery.

Recent advances in computer vision and image processing offer new opportunities for supplementing manual methods towards automation. Babawuro and Beiji (2012) detected field boundaries from satellite imagery using Canny edge detection and morphological operations, followed by a Hough transform to link the segmented boundaries. They noted that some boundaries are not well captured by their proposed method, but the exploration of integrating machine vision technology and cadastral mapping brings substantial benefits in minimizing human interventions.

When applied to very high resolution (VHR) imagery, geographic object-based image analysis (GEOBIA) is gaining popularity, which is due to the realization that image objects imply more real-world value with object appearance and topology than only pixels (Lucian Drăguţ, Tiede, & Levick, 2010). The most crucial step in OBIA is generating image objects through image segmentation. In 2018, Nyandwi applied OBIA

method in extracting cadastral parcels using chessboard and multiresolution segmentation within eCognition environment. The method was tested in one rural site and one urban site in Kigali, Rwanda, using pansharpened WorldView-2 imagery. For accuracy assessment, a tolerance of 5 meters was given to reference lines. The automation achieved 47.4% correctness and 45% completeness in rural area, while the results in urban areas were counter-intuitive. The author claimed that it was challenging for machine to extract urban parcels because the spectral reflectance for roof, garden and fences varies.

Wassie et al. (2018) investigated another image segmentation algorithm using the mean-shift segmentation plugin from QGIS to extract cadastral boundaries. They used a pansharpened and orthorectified WorldView-2 image with 0.5 m resolution. Three rural areas in Amhara Region, Ethiopia were selected as their study area, which are taken as representatives of small holder agricultural field in sub-Saharan Africa. With a 0.5-meter buffer from reference boundary, their experimental results achieved 55.4%, 28.0% and 18.9% of completeness and 16.3%, 15.9% and 7.23% of correctness for three selected sites.

Point cloud is also used for cadastral mapping. Luo et al. (2017) investigated the extraction of cadastral boundaries from airborne laser scanned data in urban areas. They designed a semi-automatic workflow including automatic extraction and post-refinement. In the automatic extraction phase, the outlines of planar object like buildings and roads are generated using Alpha shape, Canny detector, and Skeleton algorithms, while the outlines of linear object like fences are delineated by means of centreline-fitting approach. As not all extracted outlines are cadastral boundaries, manual interventions are needed in the post-refinement phase. In this phase, useful line segments are determined through visual interpretation and gaps between line segments are manually filled. Consequently, this workflow achieved an average of around 80% completeness and 60% correctness with a 4-meter tolerance from reference lines.

From the reviewed works, two main challenges for imagery-based boundary detection could be identified: 1) The complexity in extracting urban cadastral boundaries. In urban area, there are different objects within one cadastral parcel such as buildings, gardens, pavements, fences etc. which exhibit different spectral information. Moreover, the outlines of these objects contribute redundant information for cadastral boundary delineation. Therefore, it is challenging for applying image segmentation and edge detection algorithms in urban areas. 2) The need for maximizing automation accuracy and minimizing human intervention. For instance, in the work of Luo et al. (2017), the cadastral boundaries should be manually determined from all the detected buildings, roads and fences, and the gaps need to be manually filled. The above challenges provide the predominant motivation of this research. We aim at proposing a novel strategy for cadastral boundary extraction, with higher level of automation and better accuracy, as well as good performance in challenging areas such as urban environment.

2.3. Feature detection techniques

In the context of computer vision and image processing, the term feature refers to salient visual cues in digital images, such as corners, edges and blobs (Li, Wang, Tian, & Ding, 2015). Feature detection techniques including edge detection and segmentation have been applied by researchers in imagery-based cadastral mapping practices.

Segmentation is a process of dividing an image into non-overlapping segments that are homogeneous in terms of brightness, colour, texture, etc. (Pal & Pal, 1993). Supervised and unsupervised approaches are developed for image segmentation. Supervised method is often performed by training a classifier such as Convolutional Neural Networks (CNN) and Support Vector Machines (SVM). Unsupervised method defines segmentation parameters to describe characteristics like colour, texture, intensity, size and shape of image segments (Crommelinck et al., 2016). The image is then segmented based on these parameters.

Researchers claimed that segmentation-based approaches have two general drawbacks: sensitive to intra-parcel variability and dependent on parameter selection. The later often requires a prior knowledge or trial and error (García-Pedrero, Gonzalo-Martín, & Lillo-Saavedra, 2017). Multiresolution segmentation (MRS) is one of the most popular segmentation algorithm (L. Drăguţ, Csillik, Eisank, & Tiede, 2014). We choose MRS as a representative of image segmentation and compare it with our proposed method.

Classical edge detection aims to detect sharp changes in image brightness through local measurements, including first-order and second-order derivative based detection (Arbeláez et al., 2011). First-order derivative based approaches (e.g. Prewitt, Sobel) detect edges using the minimum and maximum in the first derivative with local derivative filters, while second-order derivative based approaches (e.g. Laplacian of Gaussian) using zero-crossings in the second derivative (Li et al., 2015). Derivative based edge detection is simple but noise-sensitive. Amongst others, Canny detector is justified by many researchers as a predominant one, for its better performance and capacity to reduce noise (Crommelinck et al., 2016). More recently, a remarkable progress in edge detection is the advances of learning based edge detection, which combines multiple low-level image cues into statistical learning algorithms for edge response prediction (Li et al., 2015). Globalized probability of boundary (gPb) is considered as one of the state-of-the-art methods. It involves brightness, colour and texture cues into a globalization framework using spectral clustering (Arbeláez et al., 2011). In this research, we also take gPb as a comparison algorithm against our proposed method.

In recent years, machine learning based feature detection techniques became the new trend. CNN is a remarkable representative. CNNs can learn complex hierarchical features at different layers of the network corresponding to different levels of abstraction (Bergado, Persello, & Gevaert, 2016). Traditional CNNs are usually made up of two main components, convolutional layers for extracting spatial-contextual features and fully connected feedforward networks for learning the classification rule (Bergado et al., 2016). In 2015, Long, Shelhamer, and Darrell proposed an fully convolutional network, which is adapted from contemporary CNNs. In an FCN architecture, the fully connected layers in traditional CNNs are replaced by convolutional layers. This is the reason why it is called fully convolutional networks. As compared to CNNs, the superiority of FCNs lies in the ability to perform pixel-wise prediction and accept arbitrary-sized input.

As mentioned before, FCNs can predict every pixel of the input. This is mostly realized by a downsample-then-upsample scheme (Persello & Stein, 2017). In most cases, downsampling are first adopted to capture large spatial patterns in the image, and then the coarse feature maps extracted through this process are connected back to the pixels by upsampling. SegNet is one of the famous representatives among them. It consists of an encoder network, a corresponding decoder network and a SoftMax layer (Badrinarayanan, Kendall, & Cipolla, 2017). The encoder network is used for feature extraction, including convolutional layers and max-pooling. Specifically, the convolutional layers used here is identical to the VGG 16 network (Simonyan & Zisserman, 2014). The decoder network is used to upsample the low-resolution encoder feature map to input-resolution map, including upsampling layers and deconvolutional layers. Upsampling layer uses pooling indices of the corresponding encoder max-pooling step, thus no need for learning to upsample. SegNet is proved to be efficient in terms of memory and computational time. Figure 2 shows the structure of SegNet.

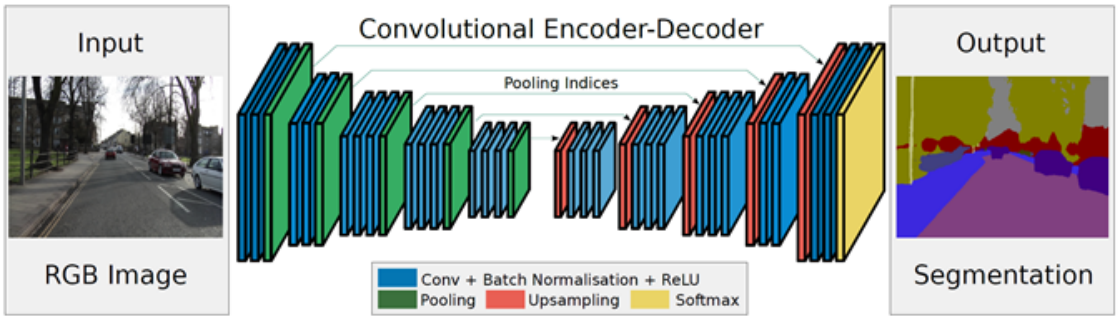


Figure 2: Architecture of SegNet (source: Badrinarayanan, Kendall and Cipolla, 2017)

In 2017, Persello and Stein proposed a novel FCN architecture (called FCN-DKs) using dilated kernels (DKs) to enlarge spatial support instead of downsampling. Correspondingly, no upsampling module is needed. The output feature map of every layer has the same spatial resolution as input. This proposed architecture is simpler and more flexible compared to existing FCNs. Table 1 shows the structure of FCN-DKs. FCN-DKs achieved very good performance in detecting informal settlements from VHR satellite images. Classification result showed that the proposed FCN outperformed state-of-the-art techniques such as Support Vector Machine (SVM) and CNNs in terms of accuracy and processing time. Inspired by their research, FCN-DKs is finally selected in this research to conduct cadastral boundary detection. Modifications are made for FCN-DKs to satisfy the demands of current study.

Table 1: Architecture of FCN-DKs

Layer	Module type	dimension	dilation	stride	pad
DK1	convolution	5×5×4×16	1	1	2
	lReLU				
	max-pool	5×5		1	2
DK2	convolution	5×5×16×32	2	1	4
	lReLU				
	max-pool	9×9		1	4
DK3	convolution	5×5×32×32	3	1	6
	lReLU				
	max-pool	13×13		1	6
DK4	convolution	5×5×32×32	4	1	8
	lReLU				
	max-pool	17×17		1	8
DK5	convolution	5×5×32×32	5	1	10
	lReLU				
	max-pool	21×21		1	10
DK6	convolution	5×5×32×32	6	1	12
	lReLU				
	max-pool	25×25		1	12
classification	convolution	1×1×32×2	1	1	0
	SoftMax				

2.4. Summary

Core concepts, prior studies and relevant feature detection techniques are reviewed in this chapter. The motivation of this research and the applied techniques are justified by basing this study on existing scientific literature. The major challenges for imagery-based boundary detection are identified, which are the difficulties in detecting urban cadastral boundaries and improving the level of automation as well as accuracy. FCN-DKs is determined to be adapted for boundary detection in our research, for its simple structure and good performance. We select MRS and gPb as representative image segmentation and edge detection algorithms to be compared with the proposed method.

3. METHODOLOGY

3.1. Study area

For the purpose of the study, two sites within Musanze district, Amajyaruguru province of Rwanda, representing urban setting and sub-urban setting respectively, were selected as case-study location. The selection is based on the availability of UAV images and the morphology of cadastral boundaries. The urban site is located in Muhoza sector and the sub-urban site is in Busogo sector. Figure 3 gives an overall view of the study area as well as the UAV images.

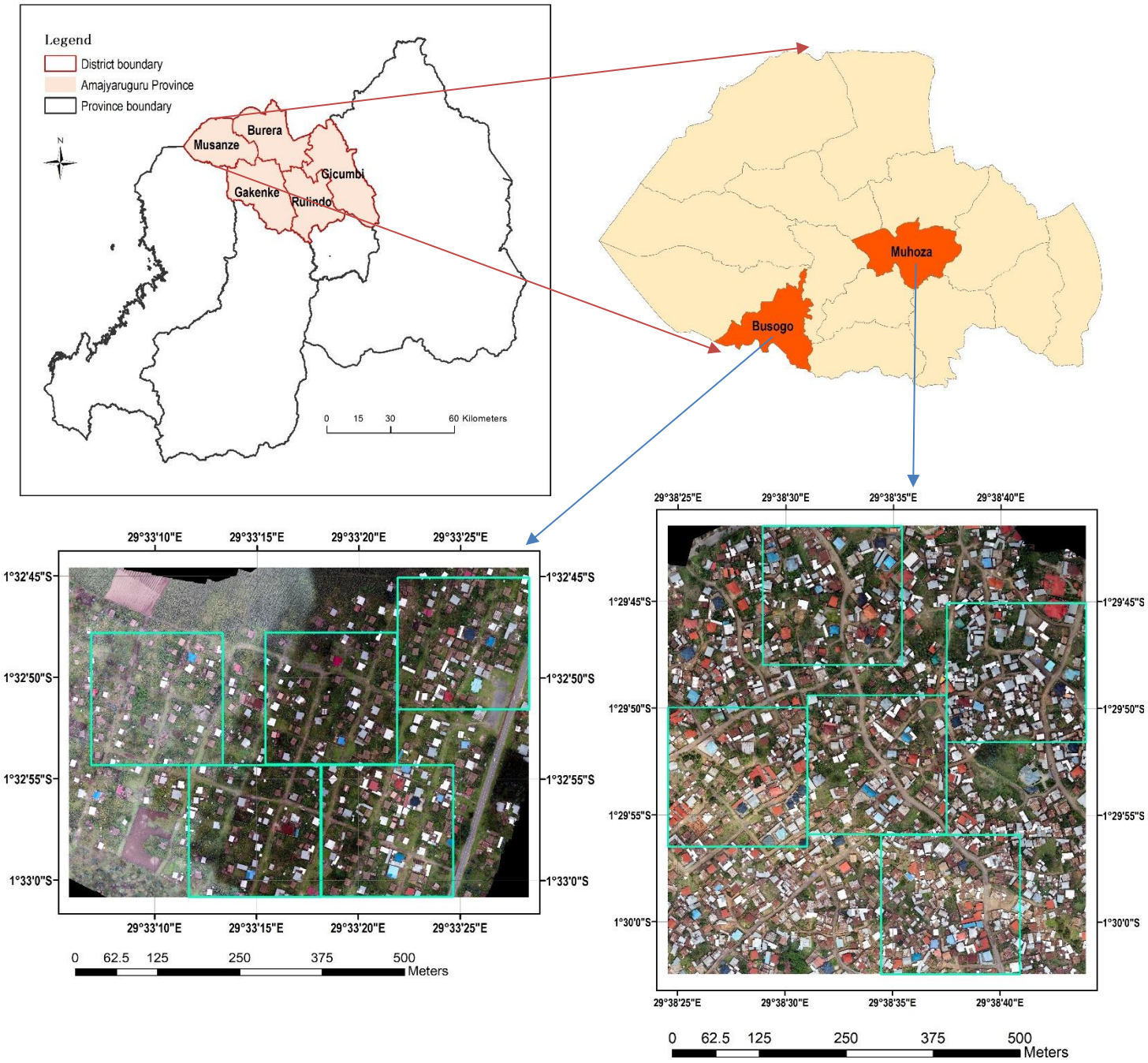


Figure 3: Study area (Busogo and Muhoza) and UAV images. Five tiles in each site (green frames) are selected to feed FCN.

In 1962, land ownership in Rwanda had changed from customary law to a system of state ownership. In 2005, a new policy was accepted called Organic Land Law (OLL) with the aim to improve land tenure security. Rwanda is one of the countries which first tested the FFP approach. Since 2008, the country has been fully covered by aerial images acquired and processed by a Dutch company (Maurice, Koeva, Gerke, Nex, & Gevaert, 2015). Even compromising with accuracy, Rwanda generated its national cadastral map based on these aerial images using participatory mapping approach. However, due to the continuously changing environment, the data is currently outdated. New technologies supporting cheap, efficient and fit-for-purpose accurate cadastral mapping will largely facilitate the data updating practices in Rwanda. Therefore, the selection of the study area has been led by the impending local demand for data update.

3.2. Overall methodology

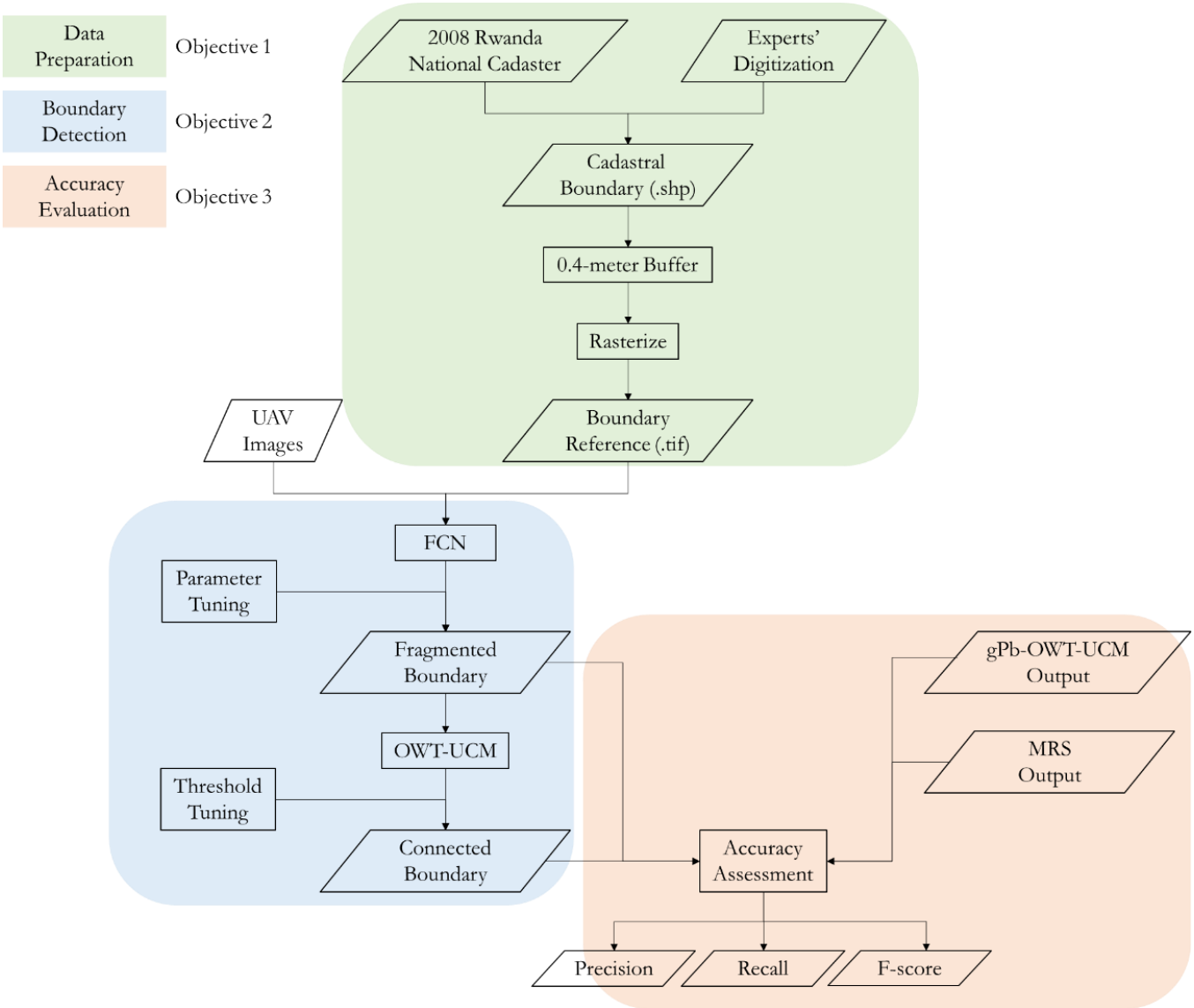


Figure 4: Overall methodology.

The flowchart above shows the overall methodology of this research, including three major parts, data preparation, boundary detection and accuracy evaluation, corresponding to the three research objectives. Boundary references are prepared with a 0.4-meter tolerance in the first part. In the second part, fragmented boundaries generated by FCN and connected boundaries produced by FCN-OWT-UCM are obtained in sequence. In this part, a systematic hyper-parameter tuning is conducted to get an optimal FCN, as well as a threshold tuning to find out the best threshold for UCM. In line with ISO standard, precision-recall framework is used for absolute accuracy evaluation in part three. We also compare the results of FCN and FCN-OWT-UCM with gPb-OWT-UCM and MRS as relative accuracy evaluation.

3.3. Data preparation

The UAV images used in this research were acquired for the its4land[1] project in Rwanda in 2018. All data collection flights were carried out by Charis UAS Ltd. These images have three bands (RGB). The spatial resolution was resampled from 0.02 m to 0.1 m considering of the balance between accuracy and computational time. Five tiles of 2000×2000 pixels were selected from each study site for the experimental analysis. Among them, three tiles were used for training (named TR1, TR2, TR3) and two for testing (named TS4, TS5) the algorithm (Figure 5&6).

For each tile, RGB layers and boundary reference were prepared as input for the classification task. The reference data was acquired by merging 2008 national cadastre and Rwandan experts' digitization. The 2008 national cadastre is outdated currently, hence experts' digitization is provided as supplements. This acquired reference was presented as polygons in a shapefile format showing the land parcels. However, to feed the FCN, the boundary reference has to be in a raster format with the same spatial resolution as RGB layers. Therefore, a series of transformation needs to be conducted. We first converted the polygons into lines of the boundaries. Rather than rasterizing the boundary lines directly, we built a 0.4 m buffer for the boundaries before rasterization. As a result, the boundary class in the reference dataset has a equal thickness of 8 pixels corresponding to 0.8 m ground sampling distance (GSD). Figure 5 and Figure 6 visualize the RGB layers and boundary reference for the selected tiles in Busogo and Muhoza.

We use the buffer with a uniform width of 0.4 m because the thickness of the boundary class directly affects the detection results. If the width of boundary class is only one single pixel, any slight displacement in detection will be regarded as a wrong detection. Therefore, a certain tolerance is often used in cadastral mapping. According to International Association of Assessing Officers (IAAO, 2015), the horizontal spatial accuracy for cadastral maps in urban environment is usually 0.3 meters or less, and in rural areas an accuracy of 2.4 meters is sufficient. Nevertheless, FFP approach advocates the flexibility in terms of accuracy to best accommodate social needs (Enemark et al., 2016). Complying with FFP land administration, we choose a 0.4-metre tolerance for urban and peri-urban environments in this research. When adopting for other applications, this number can be adjusted correspondingly and according to the demands.

[1] https://its4land.com/

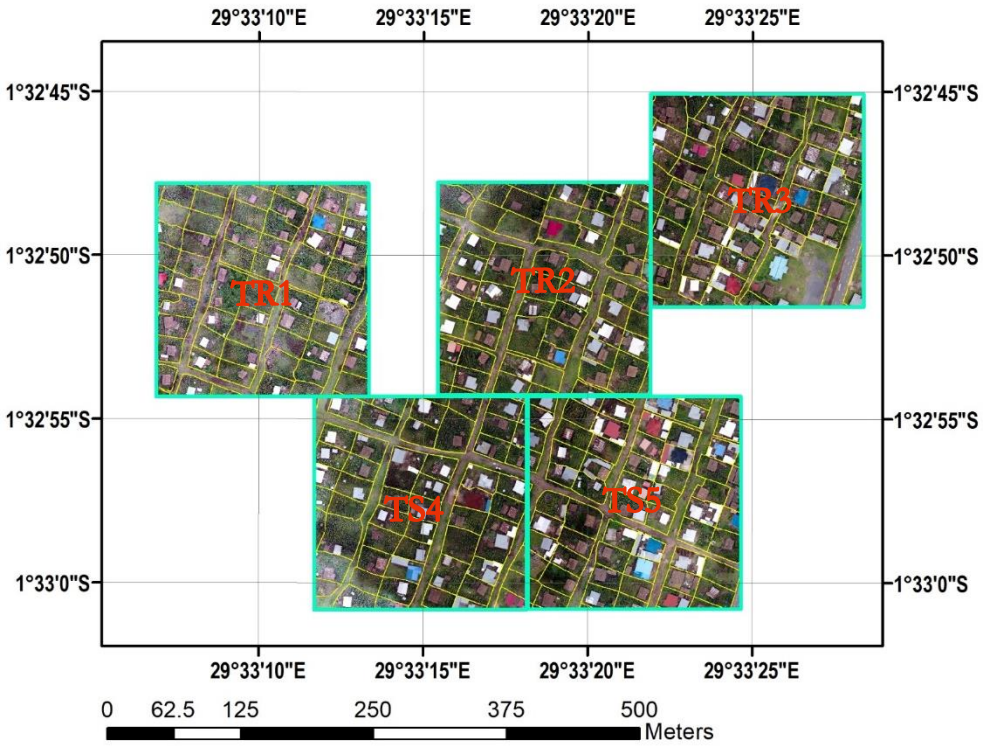


Figure 5: The UAV images and boundary reference (yellow lines) for selected tiles in Busogo.

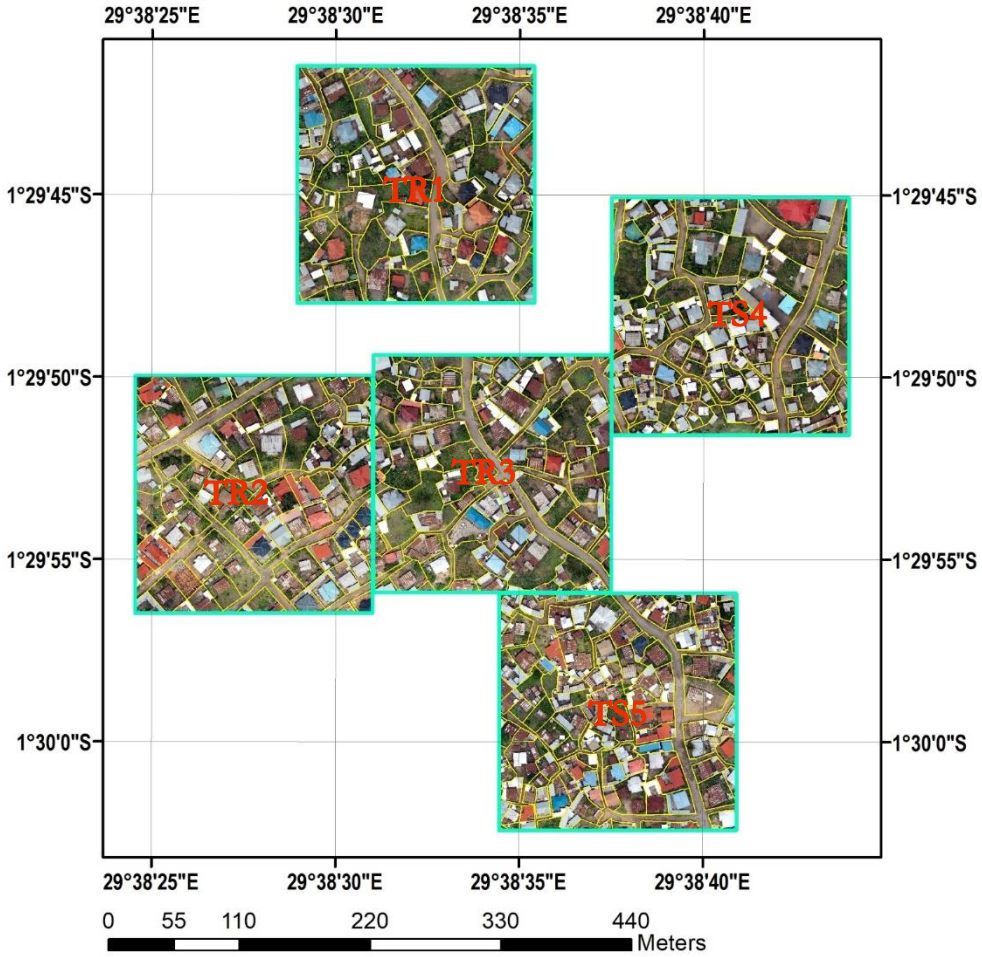


Figure 6: The UAV images and boundary reference (yellow lines) for selected tiles in Muhoza.

3.4. Automation process

3.4.1. FCN for boundary detection

The FCN used in this research is modified from the architecture of FCN-DKs as described in the paper of Persello and Stein (2017). Figure 7 shows the architecture of the proposed FCN. It consists of 15 convolutional layers interleaved by batch normalization and Leaky Rectified Linear Units (Leaky ReLU). The classification is performed by the final SoftMax layer.

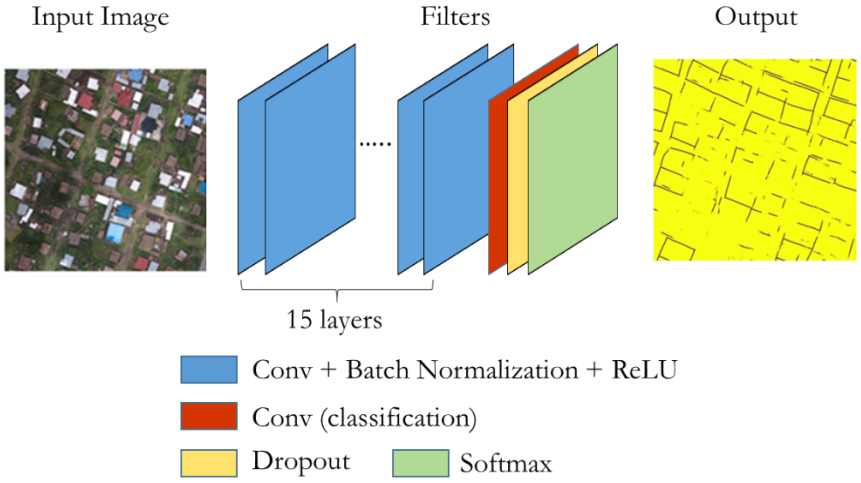


Figure 7: Architecture of the proposed FCN.

The core components of our network are the convolutional layers. Each convolutional layer consists of a bank of filters containing learnable weights and a bias term (Figure 8). The filter bank can be defined by a 4-D array $H \times W \times D \times K$, in which H, W and D namely represent the height, width and depth of one filter, and K is the number of filters. Among them, H, W and K are manual defined hyper-parameters, while D equals the number of channels of input image. Additional hyper-parameters for the convolutional layer are stride (S) and zero-padding (P). The stride is the interval by which we move the filters, and zero-padding refers to adding zeros surrounding the margins of input image before convolution. Given an input image of $M \times N \times D$, the size of output image can be expressed as following:

$$M' \times N' \times K = \left(\frac{M - H + 2P}{S} + 1\right) \times \left(\frac{N - W + 2P}{S} + 1\right) \times K \qquad (4.1)$$

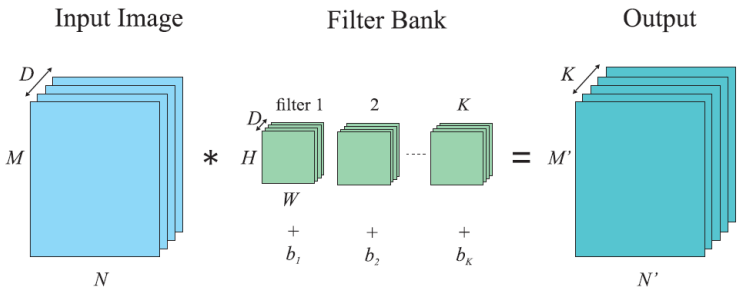


Figure 8: Convolution of an input image with a filter bank (source: Persello & Stein, 2017).

In this research, we use dilated convolution, which can capture larger receptive field without downsampling the image. This is realized by dilated kernels (DKs), which are obtained by inserting zeros between original filter elements. A dilation factor d means inserting $d-1$ zeros between each filter element. The size of the dilated kernel then becomes $H' \times W' = [d(H-1)+1] \times [d(W-1)+1]$. Therefore, the receptive field can be expanded exponentially without increasing the number of learnable parameters. In

our network, we use 3×3 kernels in the convolutional layers which are dilated increasingly from 1 to 15 to capture larger and larger contextual information. As a result, a receptive field of up to 241×241 pixels is achieved in the final layer. In each convolutional layer, zero paddings are used to keep the output feature maps at the same spatial dimension as the input. Therefore, the proposed FCN can be used to classify arbitrarily sized images directly and obtain correspondingly-sized outputs.

Batch normalization layer is used to normalize each input mini-batch. During a regular training process, the inputs to each layer are cumulatively calculated by all preceding layers. The small changes in network parameters will be magnified in deeper layers. Therefore, the distribution of each layer's input to a learning system changes as parameters of previous layers change. This phenomenon is called internal covariate shift and often requires lower learning rate and strict parameter initialization (Ioffe & Szegedy, 2015). Ioffe and Szegedy (2015) solved this problem by building batch normalization layer as part of the model architecture to normalize inputs of each layer. We took the advantage of their research and introduced batch normalization layer in our network.

Leaky ReLU is the activation function of the network. Generally speaking, the activation function is a transfer function which is added to the output of every layer. ReLU is considered as the most commonly applied activation function. It returns the value of zero given a negative input, and returns itself if the input is positive. Therefore, the function of ReLU can be written as $f(y) = max(0, y)$. However, this function suffers from the so called 'dying ReLU' problem. For negative values, the correspondent activations are 0 with constant 0 gradients, hence the weights are unable to be adjusted through gradient descent in backpropagation. Leaky ReLU fixes the problem by giving a small, non-zero gradient to negative values instead of hard-zero (He, Zhang, Ren, & Sun, 2015). The function of Leaky ReLU can be expressed as:

$$f(y) = \begin{cases} ay, & y < 0 \\ y, & y \geq 0 \end{cases} \tag{4.2}$$

Figure 9 shows the function curves of ReLU and Leaky ReLU. The latter is used in our proposed FCN.

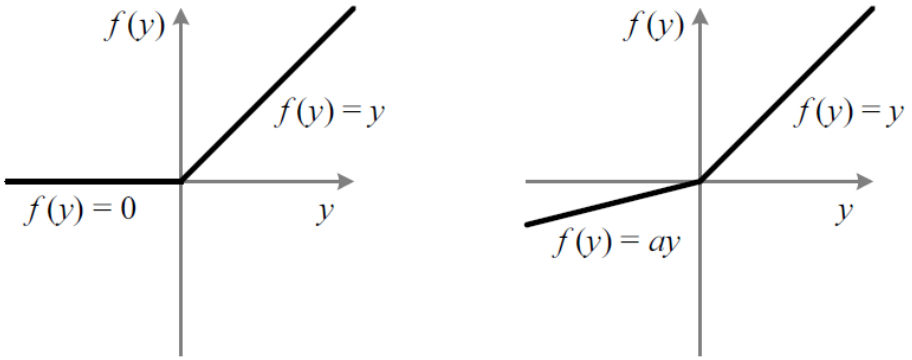


Figure 9: The function curves of ReLU (left) and Leaky ReLU (right) (source: He et al., 2015).

After constructing the framework of the FCN, hyper-parameters need to be defined. We designed a systematic tuning process to find out the optimal values for filter size, patch size, network depth, and the number of training patches. We chose one of our study areas, Muhoza, as a representative to carry out fine-tuning. The training and validation patches were generated from TR1, TR2 and TR3. TS4 and TS5 were used for testing the network performance. The hyper-parameters achieved the best classification result were selected for final implementation in both study sites. The results for parameter-tuning are stated in chapter 4.

3.4.2. Obtaining connected boundaries

FCN is trained for binary classification to distinguish between cadastral boundaries and rest on the UAV data set. Fragmented boundaries are obtained after this step. To further improve this result, the output feature map is then used to connect disjoint boundaries by applying Oriented Watershed Transform (OWT) (Arbeláez et al., 2011). The output feature map of FCN is a probability map showing the probability of each pixel being a boundary. Using the watershed transform (WT), the probability map can be converted into regions and arcs. These arcs are the potential locations of the boundaries, and the boundary strength of each arc is calculated by the average probability of all the pixels on the arc. One problem here is the probability of each pixel being a boundary varies at different orientation and the largest probability is finally assigned to each pixel as its boundary strength. Therefore, some weak arcs could be affected by its neighbouring strong arcs and end up with strong strength, as those intersecting pixels used its largest probability. OWT fixed this problem by calculating the probability of each pixel at the orientation of the arc. Therefore, the upweighted arcs could be corrected.

OWT supplies a set of closed regions surrounded by arcs with finest partition. A hierarchy of these regions can be built by the Ultrametric Contour Map (UCM) (Arbeláez et al., 2011). The hierarchy is constructed based on a graph-based region merging technique, whose merging criteria is the dissimilarity between regions. The average boundary strength of the common arcs separating two adjacent regions is regarded as the dissimilarity index. By thresholding the dissimilarity index at different scale, the most similar regions are merged iteratively, representing from over-segmentation to under-segmentation in the image.

OWT-UCM is a generic algorithm in computer vision to construct hierarchical regions from segmented contours (Arbeláez et al., 2011). As is presented in the flowchart of overall methodology (Figure 4), we adopt this technique in our research as a supplement to automatically connect the fragmented boundaries produced by FCN. We name the proposed method combining deep learning and computer vision techniques as FCN-OWT-UCM. The final boundaries are derived by thresholding the UCM. Only those arcs with a boundary strength larger than the threshold are kept.

3.5. Alternative approaches

3.5.1. Globalized probability of boundary (gPb)

Globalized probability of boundary (gPb) is proposed by Arbeláez et al. in 2011. gPb (global Pb) is a linear combination of mPb (multiscale Pb) and sPb (spectral Pb). The former conveys local multiscale Pb signals and the later introduces global information.

Multiscale Pb is an extension of the Pb detector advanced by Martin, Fowlkes and Malik (2004). The core block of Pb detector is calculating the oriented gradient signal $G(x, y, \theta)$ from the intensity images. By placing a circular disc at pixel location (x, y) and dividing it into two half-discs at angle θ, we can obtain two histograms of pixel intensity values within each half-disc. $G(x, y, \theta)$ is defined by the χ^2 distance between the two histograms. For each input image, the Pb detector divides it into four intensity images, including brightness, colour a, colour b and texture channel. The oriented gradient signals are calculated separately for each channel. Multiscale Pb modifies the Pb detector by considering the gradients at three different scales, which means we give the discs three different diameters. Therefore, we can obtain local cues at different scales, from fine to coarse structures. For each pixel, the final mPb is obtained by combining the gradients of brightness, colour a, colour b and texture channel on three scales.

Spectral Pb combines the multiscale image cues into an affinity matrix which defines the similarity between pixels. The eigenvectors of the affinity matrix which carry contour information are computed. They are treated as an image and convolved with Gaussian directional derivative filters. The sPb is calculated by combing the information from different eigenvectors.

Generally speaking, mPb detects all the edges while sPb extracts only the most salient one from the whole image (Arbeláez et al., 2011). gPb combines the two and provides uniformly better performance. After detecting the boundary probability of each pixel using gPb, we also applied OWT-UCM to the results to get connected contours. The workflow of gPb-OWT-UCM was implemented to all the testing tiles. A uniform width of 8 pixels was assigned to the detected boundary class using morphological dilation to be consistent with FCN. Accuracy assessment followed the same precision-recall framework.

3.5.2. Multiresolution segmentation (MRS)

We conducted MRS in eCognition software (version 9.4). MRS is region-merging technique starting from each pixel forming one image object or region (Baatz & Schäpe, 2000). The merging criteria is local homogeneity, which describes the similarity between adjacent image objects. The merging procedure stops when all the possible merges exceed the homogeneity criteria.

MRS relies on several parameters, which are image layer weights, scale parameter (*SP*), shape and compactness. Image layer weights define the importance of each image layer to segmentation process. In this research, we had RGB three layers in the input image. We gave them equal weights. Scale parameter is the most important parameter, which controls the average image object size (L. Drăguţ et al., 2014). A larger scale parameter allows higher spectral heterogeneity within the image objects, hence allowing more pixels within one object. Defining the proper *SP* is critical for MRS. In our research, we selected the *SP* resorting to the automatic Estimation of Scale Parameters 2 (ESP2) tool, which was advanced by L. Drăguţ et al. (2014). Shape parameter ranges from 0 to 1. It indicates a weighting between the object's shape and its spectral colour. A high value in shape parameter means less importance is put on spectral information. We set a shape parameter of 0.3 in our research. Compactness defines how compact are the segmented objects. The higher the value, the more compact the image objects may be. It was set to 0.5 in the research.

We conducted MRS to all the testing tiles and exported the output as smoothed polygons. We converted the polygons into lines in ArcGIS, giving it a 0.4 m buffer followed by a rasterization. The rasterized boundary class had 8 pixels and was compared with reference data using precision-recall measures.

3.6. Accuracy assessment

Describing the data quality in a commonly accepted way is essential for judging whether the data fulfils users' requirements or in comparison with other data set. The International Organization for Standardization (ISO 19157, 2013) provided a guideline for evaluating geographic data quality. They defined the commonly used data quality measures, including completeness, thematic accuracy, logical consistency, positional accuracy and temporal accuracy. Among them, completeness and thematic accuracy is emphasized in our research. The former is related to the presence or absence of features and the latter is about the consistency between detected boundaries and the reference data.

The results in this research are evaluated considering the classification accuracy on the testing tiles. The accuracy assessment resorted to precision-recall measures, which are a standard evaluation technique especially for boundary detection (Martin et al., 2004). Precision (*P*), also called correctness, measures the ratio of correctly detected boundary pixels to the total detected boundary pixels. Recall (*R*), also called

completeness, indicates the percentage of correctly detected boundaries to the total boundaries in reference. The F-measure (*F*) represents the harmonic mean between precision and recall (Hossin & Sulaiman, 2015). As *F* combines both precision and recall, it can be regraded as an overall quality measure. The range of these three measures is between 0 and 1. Larger values represent higher accuracy.

With the following table and formulas, how to calculate precision, recall and F-score is indicated. Pixels labelled as boundary class in both detection and reference are called True Positive (*TP*), while pixels labelled as boundary in detection but non-boundary in reference are called False Positive (*FP*). The term of False Negative (*FN*) and True Negative (*TN*) are defined similarly (Table 2). As described in section 3.3, the boundary class has a uniform width of 8 pixels. By overlaying the detection result with boundary reference, we can get the value of *TP*, *FP*, *TN* and *FN*, hence obtaining the value of precision, recall and F-score accordingly (Formula 4.3, 4.4 and 4.5).

Table 2: Confusion matrix for binary classification

	Positive Prediction	**Negative Prediction**
Positive Class	True Positive (TP)	False Negative (FN)
Negative Class	False Positive (FP)	True Negative (TN)

$$\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{4.3}$$

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{4.4}$$

$$\mathrm{F} = \frac{2 \times \mathrm{P} \times \mathrm{R}}{\mathrm{P} + \mathrm{R}} \tag{4.5}$$

3.7. Summary

This chapter describes the methods to achieve the research objectives in detail. Two study sites, Busogo and Muhoza of Rwanda are selected as representatives of sub-urban and urban settings for our research. The reference data is obtained by merging 2008 national cadastre and experts' digitization. A 0.4-metre tolerance is assigned for the reference. A workflow of FCN-OWT-UCM is designed to obtain connected boundaries. This workflow is compared against state-of-the-art techniques, gPb-OWT-UCM and MRS, on the testing tiles. The accuracy assessment goes to precision-recall measurement. Precision is a measure of correctness, corresponds to thematic accuracy in ISO 19157, while recall parallels completeness in the international standard. *F* is an overall measurement combining the two.

4. RESULTS AND ANALYSIS

4.1. Hyper-parameter optimization

The strategy used for hyper-parameter optimization is varying a single hyper-parameter while keeping the rest fixed. Muhoza is used as a representative to carry out fine-tuning. The training patches are randomly extracted from the three training tiles in Muhoza, and the average F-score of TS4 and TS5 is used to identify the optimal hyper-parameter.

Table 3 gives an overall view for the results of hyper-parameter tuning. The record of conducted tuning experiments are listed in Table 4. We searched the optimal value for filter size, patch size, the number of training patches, and network depth sequentially. Once we determine a hyper-parameter, this value will be fixed and applied in the following experiments. Stochastic gradient descent with a momentum of 0.9 was used to optimize the loss function. The training is performed in multiple stages using different learning rate. We use a learning rate of 10^{-5} for the first 180 epochs and a learning rate of 10^{-6} for another 20 epochs. The implementation of the network is based on the MatConvNet library.

Table 3: Results for hyper-parameter tuning

Hyper-parameter	Candidate Value	Optimal Value
Filter size	3×3, 5×5	3×3
Patch size	45, 85, 145, 245	245
Patch No.	900, 1500, 2100	900
Network depth	6, 10, 15	15

Table 4: Records of conducted tuning experiments

Experiment No.	Filter Size	Network Depth	Patch Size	Patch No.	Average F-score
1	5×5	3	45	900	0.25
2	3×3	6	45	900	0.34
3	3×3	6	85	900	0.38
4	3×3	6	145	900	0.41
5	3×3	6	245	900	0.42
6	3×3	6	245	1500	0.4
7	3×3	6	245	2100	0.41
8	3×3	10	245	900	0.43
9	3×3	15	245	900	0.46

4.1.1. Filter size

Experiment 1 and 2 were used to find out optimal filter size between 5×5 and 3×3. As the receptive field of two layers of 3×3 filters equals to one layer of 5×5 filter, we replaced each 5×5 convolution in experiment 1 into two layers of 3×3 convolution in experiment 2. Therefore, by comparing the result of experiment 1 and 2, the optimal filter size could be decided. With a higher average F-score, 3×3 filter outperformed 5×5 filter and was used in later experiments.

4.1.2. Patch size

The most appropriate patch size was searched by comparing experiment 2, 3, 4 and 5. In these experiments, the patch size varied from 45, 85, 145 to 245 while other hyper-parameters were kept stable. The average F-score climbed from 0.34 to 0.42 in these experiments. From Figure 10, we could notice the ascending trend of F-score by increasing patch size, and this trend is becoming mild in the later stage, closing to a saturation. The best F-score was achieved by 245×245. Therefore, patch size 245×245 was fixed as our optimal patch size.

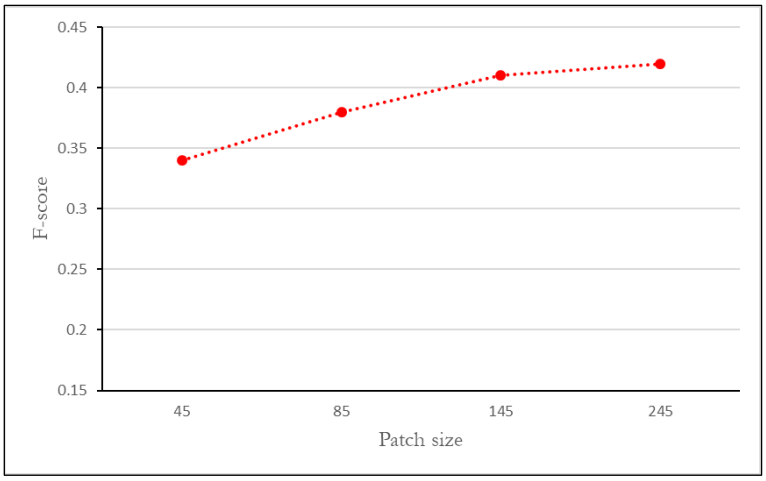


Figure 10: Average F-score of TS4 and TS5 in Muhoza with different patch size.

4.1.3. The number of training patches

In experiment 5, 6, and 7, the effect of different number of training patches to accuracy were tested. The result is quite surprising because increasing the patch number from 900 to 1500 or 2100 led to a drop in average F-score. This contradicts the expectation that larger amount of training patches would improve the accuracy. Following the experimental results, 900 training patches are used to train the network.

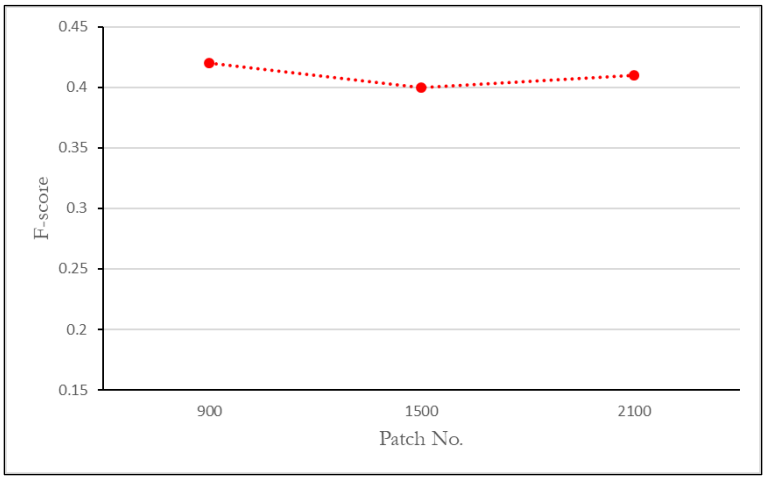


Figure 11: Average F-score of TS4 and TS5 in Muhoza with number of training patches.

4.1.4. Network depth

Experiment 5, 8 and 9 used the same filter size, patch size and patch number but different network depth. They were compared to identify the optimal network depth. According to Figure 12, larger F-score was obtained with deeper network. A network depth of 15 convolutional layers is employed for the final implementation.

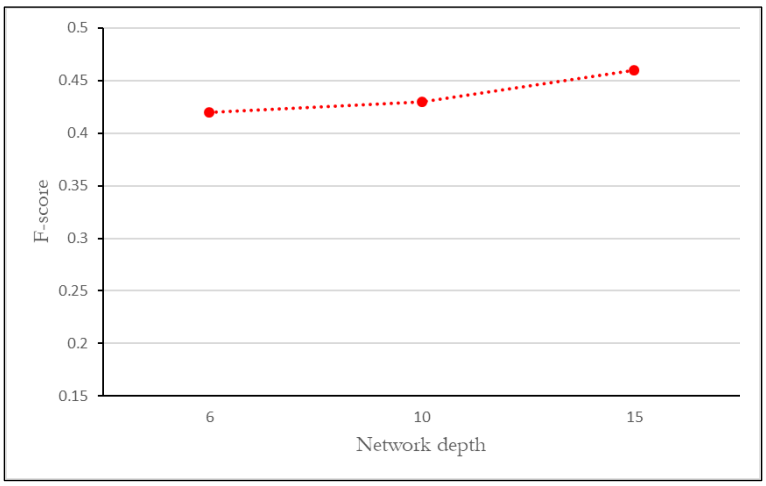


Figure 12: Average F-score of TS4 and TS5 in Muhoza with different network depth.

4.2. Threshold tuning

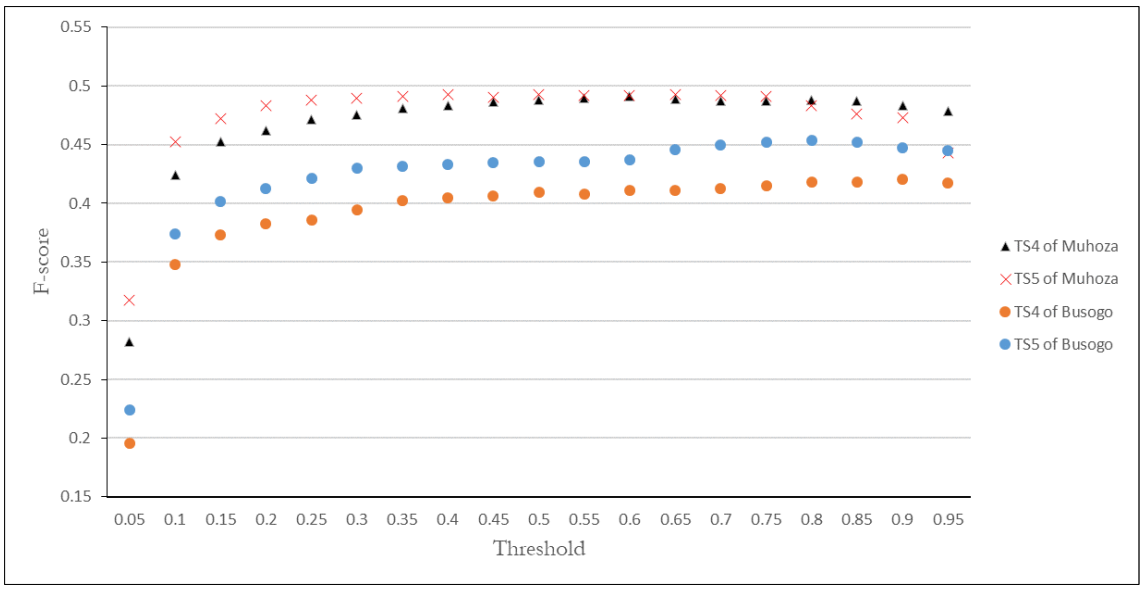


Figure 13: Threshold tuning result for the resting tiles in Busogo and Muhoza.

As mentioned in section 3.4.2, the real-valued image of a UCM is obtained by setting a certain threshold. FCN-OWT-UCM results in different contours with different thresholds. Therefore, we test every tile for 20 thresholds by increasing its value from 0 to 1 with an interval of 0.05. The optimal result with the highest F-score is taken as the final result of FCN-OWT-UCM, and the corresponding threshold is designated. The experimental results of threshold tuning for the testing tiles are shown in Figure 13. According to the result, in Busogo, the optimal performance was achieved with a threshold of 0.9 for TS4 and 0.8 for TS5; while in Muhoza, the best threshold for TS4 and TS5 is 0.6 and 0.65 respectively. Interestingly, a common trend could be witnessed for these testing tiles in Figure 13. In the early stages of

threshold tuning, the F-score is ascending steadily. When the threshold kept increasing from around 0.3, the F-score remains relatively stable. The effect of changing threshold to F-score is not so obvious from that point onwards.

Figure 14 visualizes the output boundary map under different thresholds, taking TS5 of Muhoza as an example. By varying the threshold from 0.05, 0.25, 0.65 to 0.95, the boundary pixels in the output map is decreasing continuously. The output map with a threshold of 0.05 is rather messed up. In contrast, a threshold of 0.95 generates a very clean output showing only the strongest boundaries. Stuck in the middle, a threshold of 0.65 derives the optimal result.

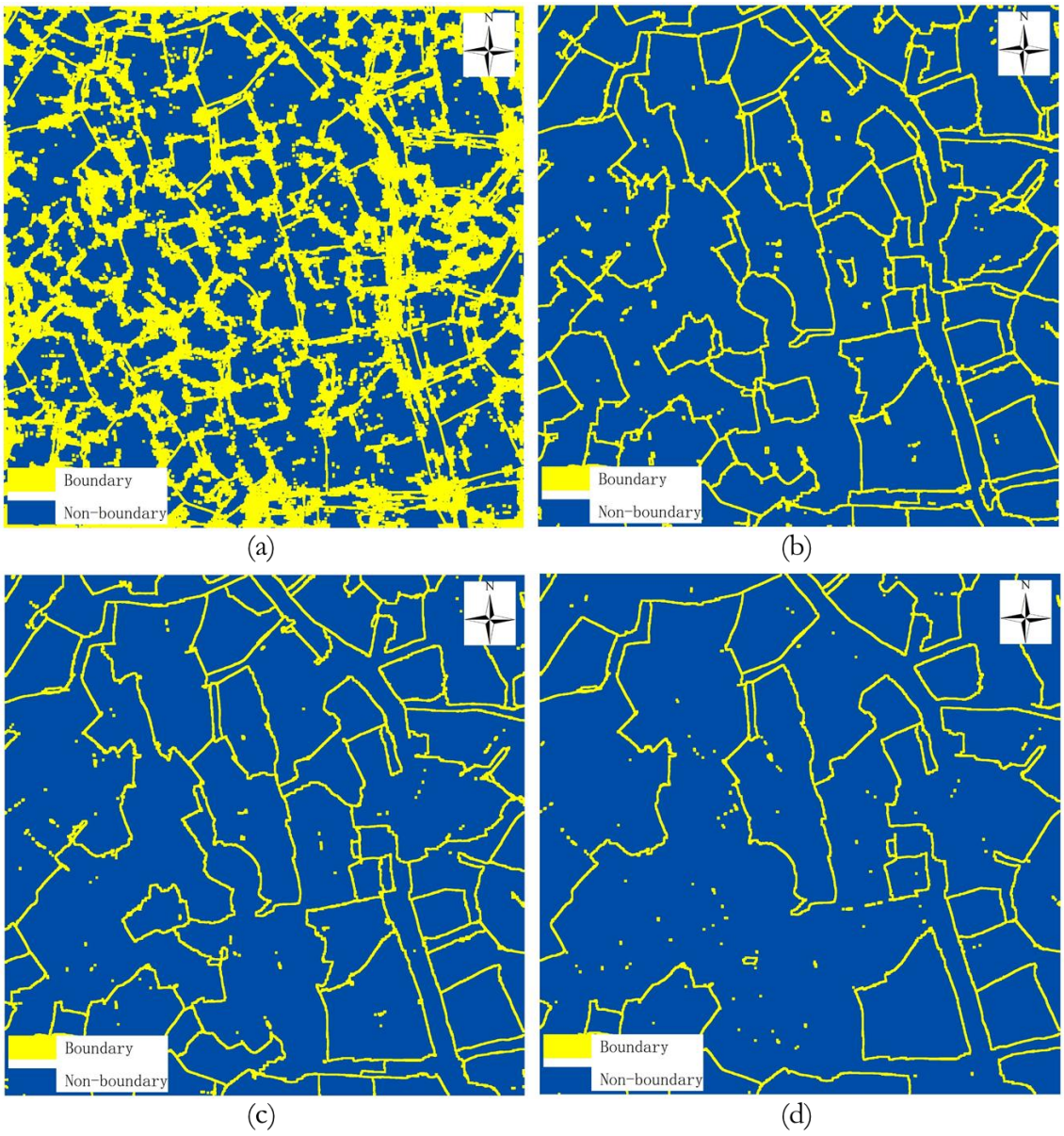


Figure 14: The output boundary map of TS5 in Muhoza using different thresholds. The thresholds for (a), (b), (c), (d) are namely 0.05, 0.25, 0.65, 0.95.

4.3. Final implementation on both study sites

The experimental results include two parts, fragmented boundaries generated by FCN and connected boundaries produced by FCN-OWT-UCM. The proposed method is implemented on both study sites, Busogo and Muhoza, to test its generalization ability. As mentioned in chapter 3, the performance of the proposed method is evaluated based on the testing tiles using precision-recall framework. Two testing tiles on each site are selected, in order to obtain an unbiased estimation. The visual and numerical results of the testing tiles are demonstrated in the following figures and tables. To see the results of the training tiles please refer to the appendix 1.

Table 5: Precision, recall, and F-score for testing tiles using FCN and FCN-OWT-UCM in Busogo.

Method	Tile	*P*	*R*	*F*
FCN	TS4	0.54	0.34	0.42
FCN-OWT-UCM		0.44	0.41	0.42
FCN	TS5	0.62	0.35	0.45
FCN-OWT-UCM		0.46	0.45	0.45

Table 6: Precision, recall, and F-score for testing tiles using FCN and FCN-OWT-UCM in Muhoza.

Method	Tile	*P*	*R*	*F*
FCN	TS4	0.61	0.38	0.47
FCN-OWT-UCM		0.45	0.54	0.49
FCN	TS5	0.66	0.33	0.44
FCN-OWT-UCM		0.48	0.50	0.49

Table 5 presents the classification accuracy of Busogo. The classification result of FCN in TS4 gets 0.54 in precision, which means the ratio of true boundaries to the total detected boundaries is 54%. The value of recall is 0.34, indicating 34% of cadastral boundaries among all the boundaries in reference are detected. Other results from Table 5 and Table 6 could be comprehended in the same way. Similarly, in TS5, 35% of cadastral boundaries are detected by FCN, and the correctness of detection is 62%. TS4 has an F-score of 0.42 and TS5 has an F-score of 0.45. FCN-OWT-UCM achieved different precision and recall as FCN, while they ended up with the same F-score in both tiles.

Table 6 shows the classification result of Muhoza. In TS4, a precision (or correctness) of 0.61, a recall (or completeness) of 0.38 and a F-score of 0.47 are achieved by FCN. And there is a slight increase in F-score after applying FCN-OWT-UCM, raising to 0.49. Similar trend can also be witnessed in TS5. From FCN to FCN-OWT-UCM, the F-score increased from 0.44 to 0.49.

The average results of two testing tiles in each study area using FCN and FCN-OUT-UCM are calculated respectively, to compare the results of the two study sites as well as the performance of FCN and FCN-OWT-UCM. These results are visualized in Figure 15 and Figure 16. Comparing these two figures, the results of both FCN and FCN-OWT-UCM in all the metrics get a higher value in Muhoza than in Busogo. Moreover, in both figures, FCN-OWT-UCM manifests lower precision and higher recall than FCN. In Busogo, FCN-OUT-UCM has the same F-score as FCN, whereas in Muhoza, FCN-OWT-UCM outperforms FCN.

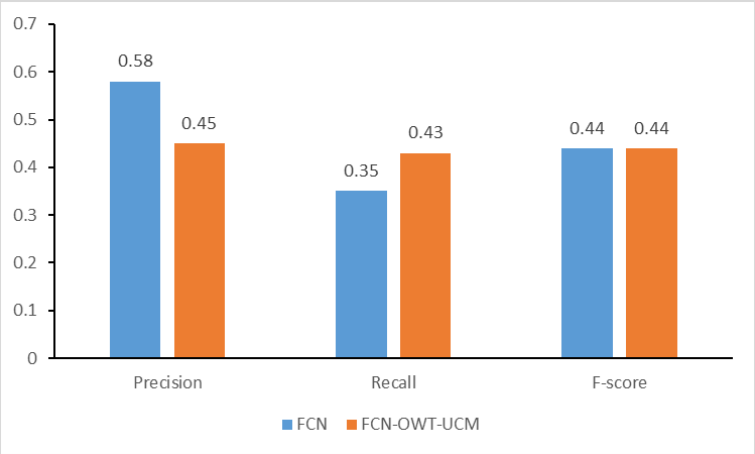


Figure 16: Average results of two testing tiles using FCN and FCN-OWT-UCM in Busogo.

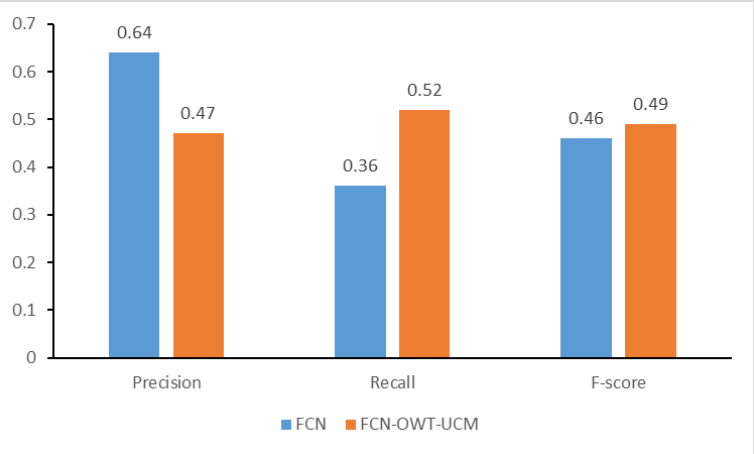


Figure 15: Average results of two testing tiles using FCN and FCN-OWT-UCM in Muhoza.

The output feature maps of FCN are presented in Figure 17. As explained in section 3.4.2, the feature map is a probability map showing the probabilities of each pixel being a boundary, and is used as input for generating connected boundaries by applying OWT-UCM. The feature maps displayed in Figure 17 are polarized with the majority of the pixels showing a very high (yellow pixels in the image) or very low (dark blue pixels in the image) boundary probability. Only a few pixels possessing moderate boundary probability (light blue pixels).

Figure 18 and 19 visualize the final output maps of FCN and FCN-OWT-UCM. Comparing with Figure 17, the feature maps of FCN are very close to the final output of FCN. Both FCN and FCN-OWT-UCM delineates very clean and clear boundaries. Although building outlines correspond to strong edges, they are not confused by the FCN with cadastral boundaries. However, there are still several missing fragments of the boundaries. These missing boundaries are mainly invisible boundaries. Besides, some boundaries marked by roads or building walls are also not detected by FCN. In general, most detected boundaries coincide with fences and hedges. FCN-OWT-UCM closed some land parcels correctly, but also brought in some false detections.

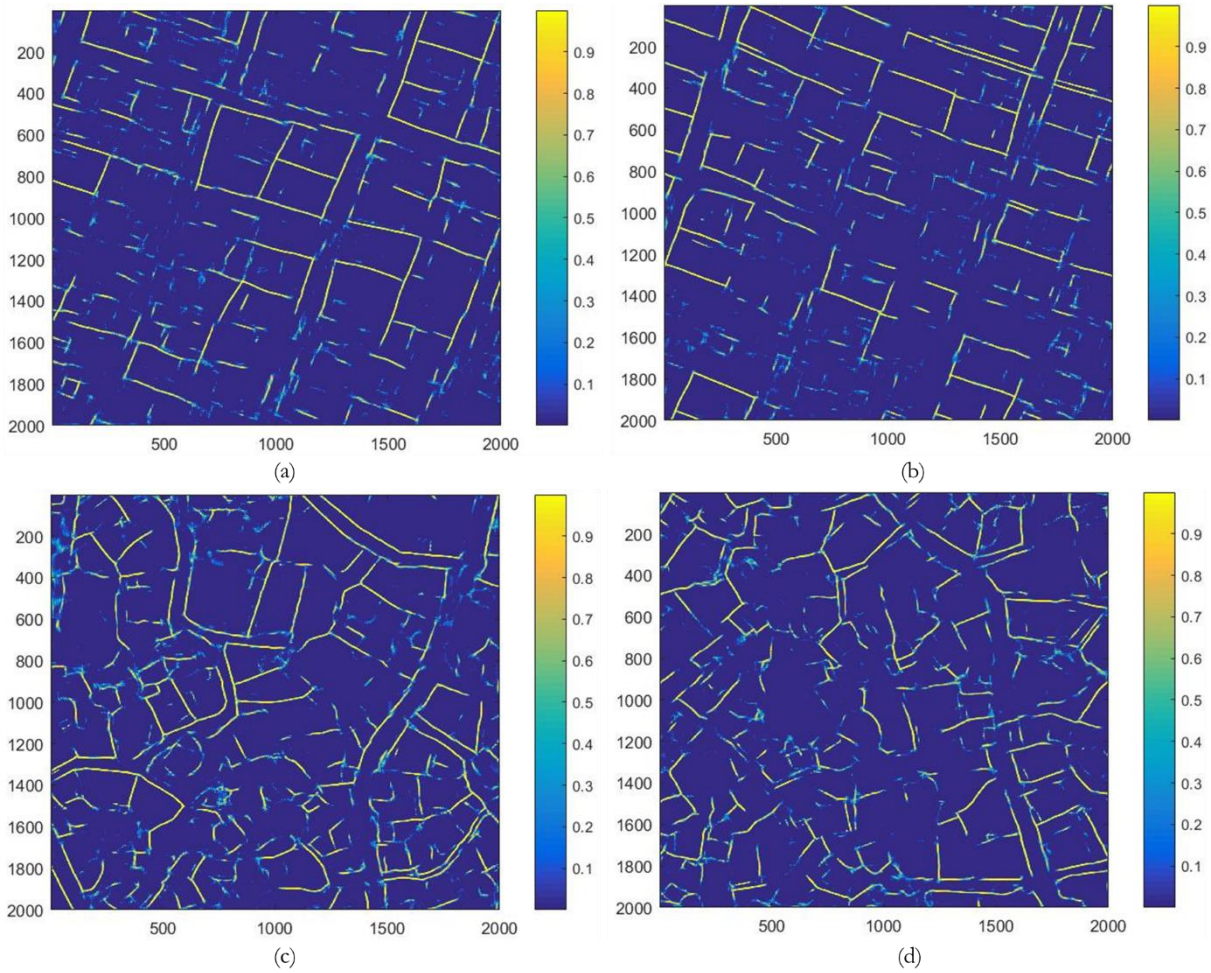


Figure 17: Output feature maps of FCN. (a), (b) are the feature maps of TS4 and TS5 in Busogo; (c), (d) are the feature maps of TS4 and TS5 in Muhoza.

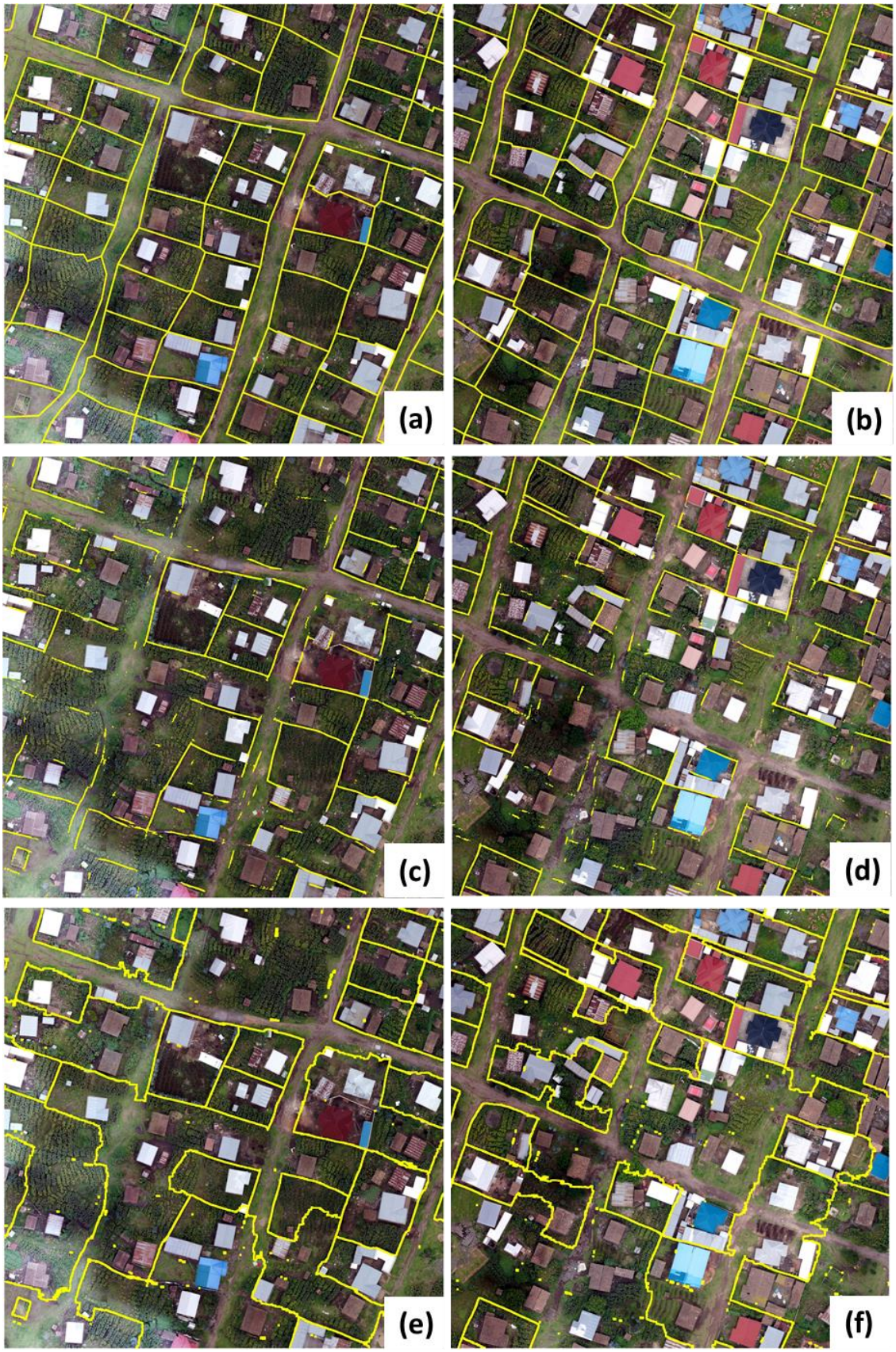


Figure 18: Results of TS4 (a, c, e) and TS5 (b, d, f) in Busogo. (a, b) are boundary references; (c, d) are classified maps of FCN; (e, f) are classified maps of FCN-OWT-UCM.

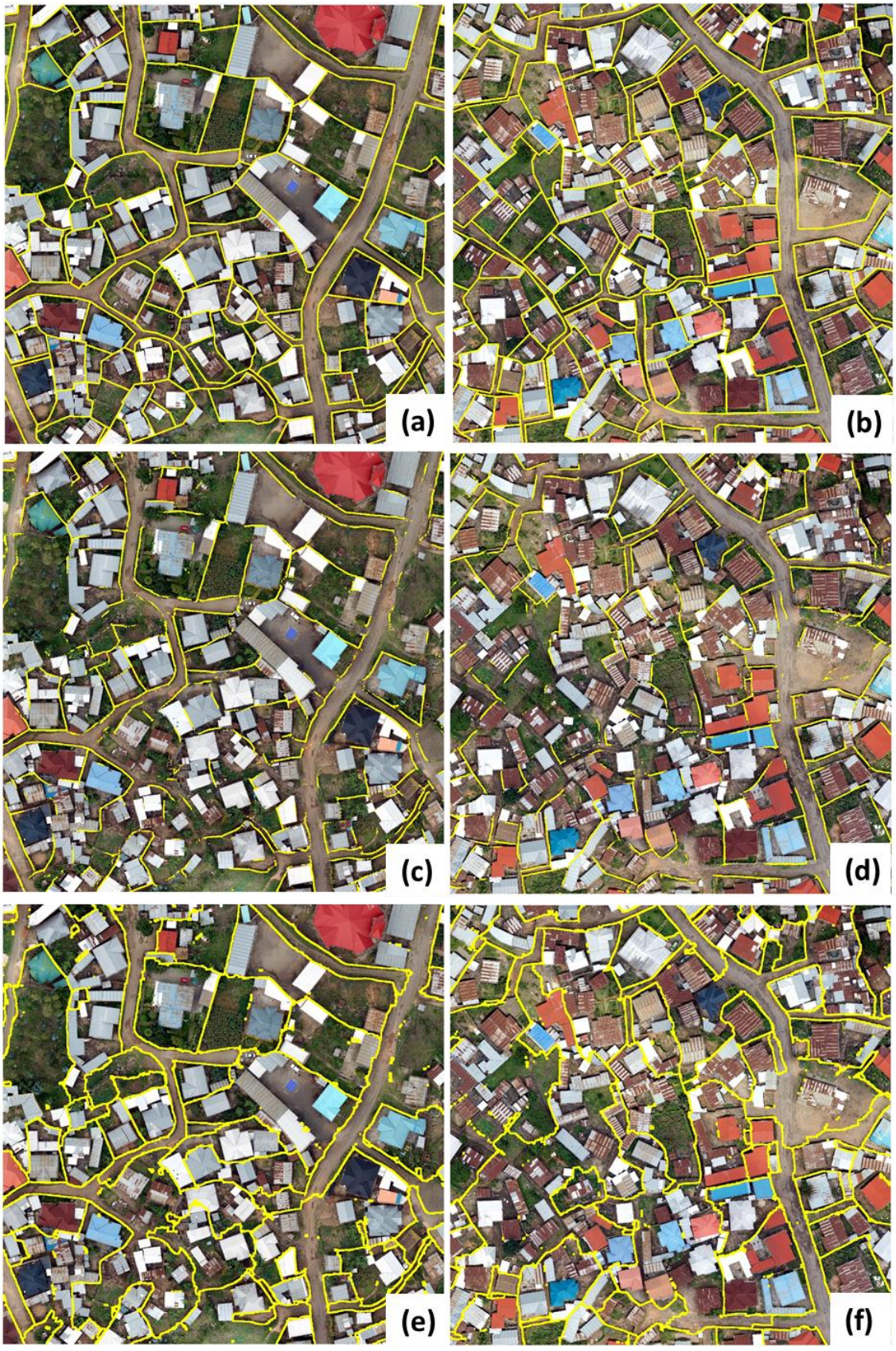


Figure 19: Results of TS4 (a, c, e) and TS5 (b, d, f) in Muhoza. (a, b) are boundary references; (c, d) are classified maps of FCN; (e, f) are classified maps of FCN-OWT-UCM.

4.4. Alternative approaches

4.4.1. Globalized probability of boundary (gPb)

As mentioned in section 3.5.1, the output of gPb is a boundary probability map. We used OWT-UCM to convert the boundary probability map into hierarchical regions. The closed contours of these regions can be regarded as the detected cadastral boundaries. The optimal threshold for UCM was tuned 20 times for each tile following the same method as described in section 4.2. The records of threshold tuning are attached in appendix 2. To make a fair comparison with other methods, we dilated the width of output boundary class from one pixel to 8 pixels using morphological dilation, which equals the width of boundary class in reference. Figure 20 and Figure 21 shows the output of gPb-OWT-UCM on the testing tiles of Busogo and Muhoza. Accuracy assessment results of these tiles are presented in table 7, using precision-recall measures.

Figure 20: The detected boundaries of TS4 (c) and TS5 (d) in Busogo using gPb-OWT-UCM. (a, b) are boundary references. The threshold of UCM is 0.15 in both (c) and (d).

Figure 21: The detected boundaries of TS4 (c) and TS5 (d) in Muhoza using gPb-OWT-UCM. (a, b) are boundary references. The threshold of UCM is 0.3 in (c) and 0.1 in (d).

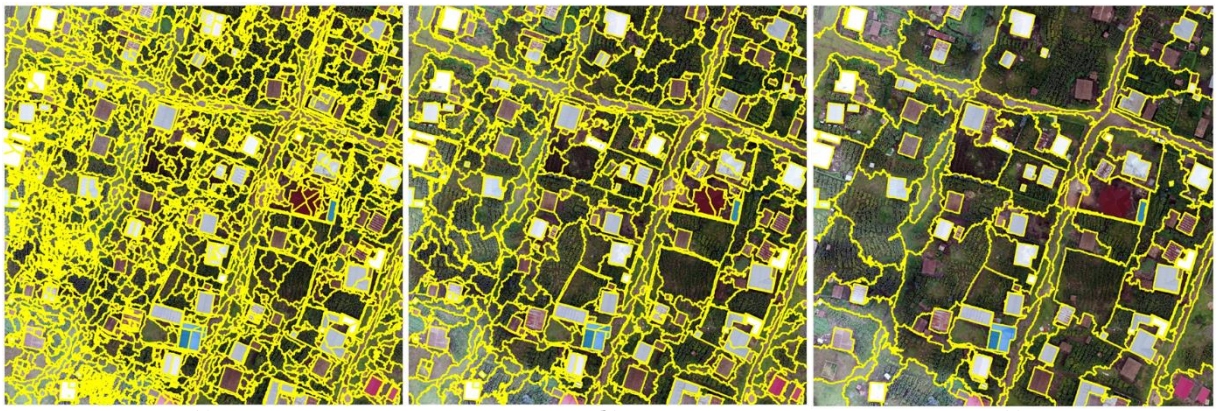
Table 7: Precision, recall and F-score for testing tiles in Busogo and Muhoza using gPb-OWT-UCM.

Location	**Tile**	***P***	***R***	***F***
Busogo	TS4	0.20	0.55	0.30
	TS5	0.22	0.68	0.33
Muhoza	TS4	0.18	0.41	0.25
	TS5	0.13	0.51	0.21

From table 7, gPb-OWT-UCM obtained an average F-score of 0.32 in Busogo and 0.23 in Muhoza. The performance of gPb detector is better in Busogo than in Muhoza. Moreover, the results of gPb detector in both study sites held a relatively high value in recall but low in precision. Associated with figures above, gPb shows a very strong ability in contour detection. Almost all the contours in the images were detected by gPb, including fences, building outlines and even roads which were not well detected by FCN. However, not all the contours are cadastral boundaries. gPb detector lacks the ability to abstract cadastral boundaries from the detected contours. Therefore, it obtained a high recall but very low precision, resulting in a relatively poor F-score.

4.4.2. Multiresolution segmentation (MRS)

As stated in section 3.5.2, the key control of MRS is the scale parameter (*SP*). We applied the ESP2 tool to identify the proper *SP*. It produces fully automated three-scale level segmentations based on the local variance. The authors claimed that image objects of different sizes are best presented at different scales, hence multi-scale segmentation is more suitable than single-scale to model image objects in a scene (L. Drăguţ et al., 2014). This tool extracts 3 scale levels, representing from finest to coarser segmentation. Each scale parameter can be generated independently, based on the pixel level, or within a hierarchy, where a parent-child relationship exists between the levels. In our research, we selected the non-hierarchy version to capture cadastral parcels based on pixel level, as the intraclass variance is relatively large.



(a) (b) (c)

Figure 22: The three-scale segmentations produced by ESP2 on TS4 in Busogo. (a) is level 1 with a *SP* of 59; (b) is level 2 with a *SP* of 111; (c) is level 3 with a *SP* of 301.

Table 8: MRS results on scale level 2 and level 3 for all the testing tiles in Busogo and Muhoza. For each tile, the level marked by red colour is the finally selected scale level for MRS.

Location	Tile	Level	*SP*	*P*	*R*	*F*
Busogo	TS4	2	111	0.12	0.51	0.19
		3	301	0.14	0.32	0.20
	TS5	2	151	0.16	0.55	0.24
		3	401	0.19	0.39	0.25
Muhoza	TS4	2	251	0.18	0.54	0.27
		3	801	0.15	0.19	0.17
	TS5	2	211	0.22	0.66	0.33
		3	1201	0.27	0.12	0.17

Figure 22 presents the three scale levels produced by ESP2 on TS4 in Busogo. The scale parameter for each level are selected automatically by ESP2. Level 1 is the finest segmentation, where the image objects are too small as compared to cadastral parcels. Therefore, for each testing tile, we selected level 2 and level 3 for accuracy assessment, using the scale with highest accuracy as the final segmentation. Table 8 shows results of accuracy assessment on all the testing tiles. According to table 8, we selected the segmentation of level 3 for TS4 and TS5 in Busogo, and level 2 for TS4 and TS5 in Muhoza. The selection was based on a higher F-score. Numerical results indicated that MRS had better performance in Muhoza than in Busogo. The former achieved an average F-score of 0.3, and 0.23 for the latter. Similar to gPb detector, MRS had high recall value while low precision value. The final results of MRS are visualized in Figure 23 and Figure 24. As in the case of gPb detector, buildings remain an obstruction for MRS to detect homogeneous cadastral parcels.

Figure 23: Results for MRS on TS4 (c) and TS5 (d) in Busogo. (a, b) are boundary references. The value of *SP* is 301 in (c) and 401 in (d).

Figure 24: Results for MRS on TS4 (c) and TS5 (d) in Muhoza. (a, b) are boundary references. The value of *SP* is 251 in (c) and 211 in (d).

4.5. Performance comparison

To get an unbiased assessment, we evaluate the performance of each method based on the average score of the two testing tiles in each study site. The results are displayed in Figure 25 and Figure 26. FCN-OWT-UCM held the highest F-score in both study sites, closely followed by FCN. gPb-OWT-UCM and MRS took the third and fourth place in Busogo, whereas in Muhoza, their ranking was reversed. In both study sites, FCN outperformed other algorithms in terms of precision. The highest recall was achieved by gPb-OWT-UCM in Busogo and MRS in Muhoza. However, these two algorithms had very poor performance in terms of precision. FCN-OWT-UCM had a relatively balanced precision and recall, leading to the best overall performance.

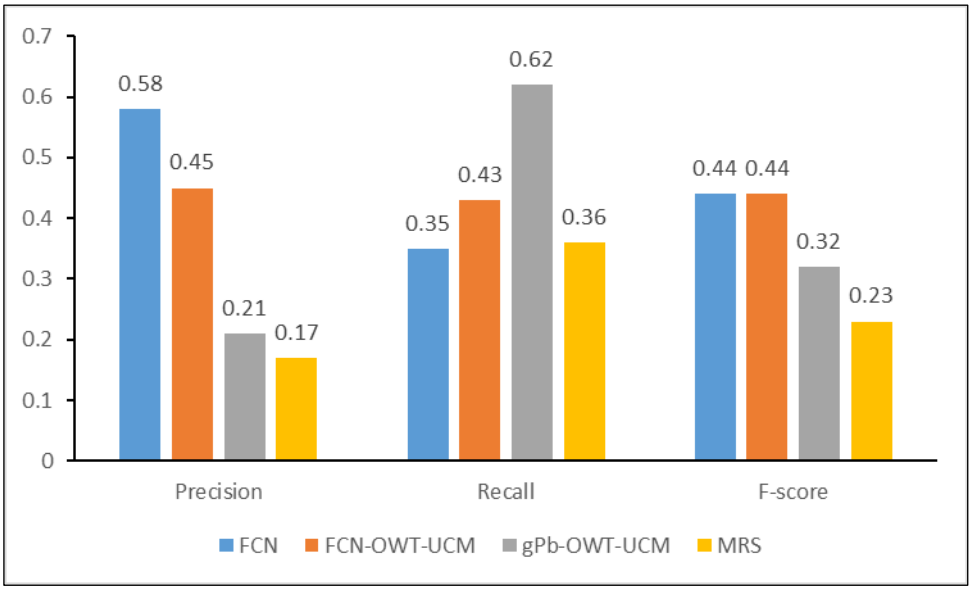


Figure 25: The performance of FCN, FCN-OWT-UCM, gPb-OWT-UCM and MRS in Busogo.

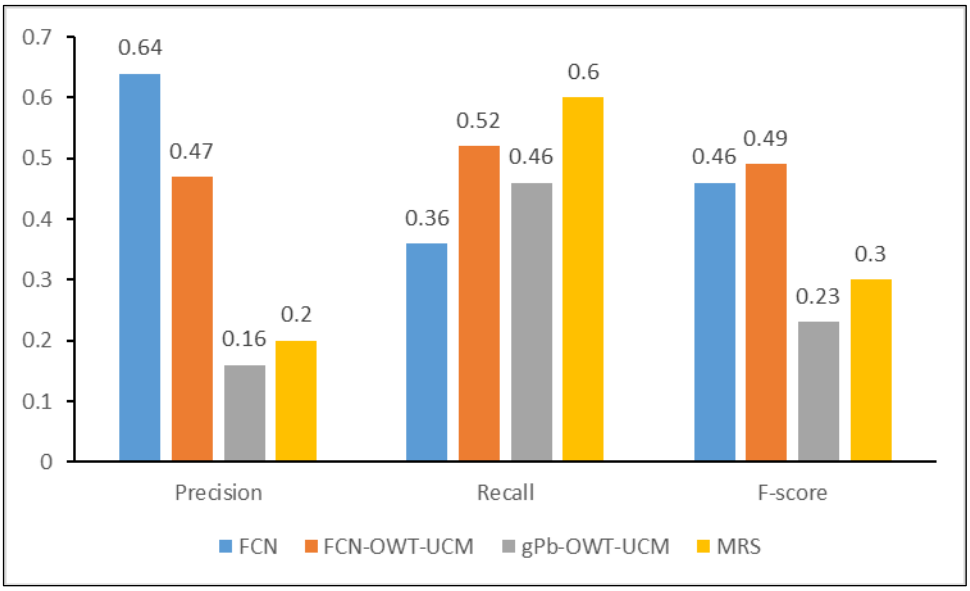


Figure 26: The performance of FCN, FCN-OWT-UCM, gPb-OWT-UCM and MRS in Muhoza.

4.6. Summary

In this chapter, the results for hyper-parameter tuning, threshold tuning and the final implementation are thoroughly presented and described. The final implementation is conducted using 15 layers of 3×3 filter with increasing dilation factor from 1 to 15. 900 training patches (300 from each training tile) with patch dimension 245×245 are used to train the FCN. Final outputs of FCN-OWT-UCM is obtained by thresholding the Ultrametric Contour Map 20 times to select the optimal result. We present the results of the final implementation in two parts, fragmented boundaries of FCN and connected boundaries of FCN-OWT-UCM. These results are then compared with gPb-OWT-UCM and MRS. From the results, we have the following findings:

a) The detection results of Muhoza is better than Busogo using the proposed method.
b) FCN-OWT-UCM can obtain slightly better or at least equal results as FCN.
c) The output feature maps of FCN are polarized and visually similar to the final output of FCN.

d) Most detected boundaries of FCN coincide with fences or hedges.
e) FCN-OWT-UCM can connect the fragmented boundaries but also introduces false prediction.
f) gPb-OWT-UCM and MRS can obtain better recall than FCN and FCN-OWT-UCM, but they have very poor performance in precision. As a result, FCN and FCN-OWT-UCM achieve the best overall performance in cadastral boundary detection.

5. DISCUSSION

Considering the constraints of the invisible cadastral boundaries, land administration professionals perceived that a 40 to 50 percent of automatic delineation would be very significant in reducing time and labour involved in cadastral mapping practices (Wassie et al., 2018). This goal is reached by the proposed method with a 0.4-metre tolerance in both Busogo and Muhoza, indicating good generalization and transferability of the proposed approach in cadastral boundary mapping.

We tested the effects of different hyper-parameters of FCN to classification accuracy as a starting point for boundary detection. Experiments on filter size showed that two 3×3 filters performed better than one 5×5 filter. With the same receptive field, two 3×3 filters have only 18 parameters while one 5×5 filter has 25 parameters. What's more, compared to single larger-sized filter, multiple small filters are interleaved by activation functions, resulting in better abstraction ability. Therefore, with less learnable parameters and better feature abstraction ability, smaller filters are preferred.

Experiments on patch size demonstrated a first increase then closing to saturation trend. Larger patch size stands for lager contextual information, which explains for the ascending trend. However, with the limitation of the receptive field, when the patch size is increased further, far beyond the receptive field, the contextual information is saturated to the network.

We also studied the effects of training sample size, however we obtained surprising result. Generally speaking, large training samples would improve the performance as well as generalization ability of FCN. However, in our experiments, increasing the number of training patches led to a slight decrease in accuracy. Considering the fact that we use very large patch size with dimension 245×245, and the size of the training tiles is 2000×2000. With 16 unoverlapped patches, we can cover the whole image. In our experiment, we randomly extracted 300 patches from each training tile (900 in total). There were already many overlapped areas. Further increasing the number of training samples does not necessarily add in new information. Therefore, increasing training samples did not improve the accuracy in our context.

We used a network depth of 15 convolutional layers. Deeper network corresponds to larger receptive field and higher level of abstraction. Hierarchical features could be learnt through different layers of the network. Furthermore, we used an increasing dilated factor in each convolutional layer, a final receptive field of 241×241 was achieved. With a deep network, larger areas in the images could be scanned by the network, which means larger patterns could be extracted.

After setting the proper hyper-parameters for FCN, the thresholds for UCM were also tuned. In this research, we tested 20 thresholds for UCM and selected the optimal result by validating with reference. However, in real classification tasks, there is no reference data available. We can only decide the threshold via visual perception. According to the experimental results, by increasing the thresholds, the F-score increases first and tends to be stable with slight fluctuation after a certain point. Therefore, in real practices, although influenced by human subjectivity, threshold selection will not largely affect the output.

In the final implementation, results showing that the true positive detected by FCN are mainly fences and hedges. This can explain why the classification result in Muhoza is slightly better than Busogo, considering the fact Muhoza has denser buildings and fences. The output of FCN-OWT-UCM is only a slightly better or even equal to FCN. This is related to the output feature map of FCN, which is the input for OWT-UCM. The feature map is polarized with mainly strong boundaries or non-boundaries. Very few weak

boundaries with moderate boundary probability exist in the feature map. Therefore, there are not so many cues for OWT to connect the fragmented boundaries correctly, resulting in just very little improvement.

Compared with alternative edge detection and image segmentation approaches, FCN and FCN-OWT-UCM achieved better overall performance. Lacking abstraction ability, standard edge detection and image segmentation cannot fill the semantic gap between high-level cadastral boundary concept and low-level image features. As a result, gPb detector and MRS achieved high recall but low precision. In other words, gPb cannot determine cadastral boundaries from all the detected contours, while MRS cannot eliminate over-segmentation caused by the spectral differences within one cadastral parcel.

FCN can supply high precision, while gPb and MRS can supply high recall. Therefore, for further study, we can consider a combination of these methods. We can combine them in two ways. The first one is to involve the output of gPb or MRS along with UAV images as input for FCN. FCN has strong feature learning ability. It is possible that FCN can determine cadastral boundaries from the output of gPb or MRS, hence increasing both precision and recall. The second way is to linearly combine the feature map of FCN and the probability map of gPb, followed by applying OWT-UCM to the output. Therefore, compared to use only the feature map of FCN, OWT have additional cues from gPb to connect the cadastral boundaries correctly, hence increasing the accuracy.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Refelection to research objectives and questions

The purpose of this research is to propose a novel machine-based method which can facilitate cadastral mapping and data upgrading practices by reducing time, cost and human intervention. This generic objective is divided into three specific objectives and several research questions, which form the sketch of the whole research. These questions are answered via our study and they are concluded as follow:

Objective 1: To prepare the data for cadastral boundary detection.

a) How to prepare the reference data?
b) How to design the training and testing dataset?

Data preparation is the first step. Our study areas locate in two sectors of Rwanda, Busogo and Muhoza. The reference data and the UAV images should be prepared as input for FCNs. Reference is adapted from 2008 Rwanda cadastre and experts' digitization, using a 0.4 m tolerance. Three tiles of training data and two tiles of testing data are selected from each site. The tile size is 2000×2000 and the spatial resolution for reference and input images is 0.1 m.

Objective 2: To develop a methodology for cadastral boundary detection.

a) Which FCN architecture is appropriate for boundary detection?
b) What are the optimal hyper-parameters for the proposed FCN?
c) How to improve the result of FCN?

Boundary detection is the core component of the study. We adjust the architecture of FCN-DKs proposed by Persello and Stein (2017) for our research and optimised the hyper-parameters through a systematic tuning. The result of FCN is fragmented boundaries, which can be improved by using OWT-UCM to get connected.

Objective 3: To compare the performance of the proposed method with other state-of-the-art methods.

a) What are the widely accepted metrics for evaluating boundary detection?
b) Which method achieves the best result?

The last step is accuracy assessment. We checked the ISO standard for describing geo-spatial data quality and selected metrics that can be adopted in our context. We determined to use precision-recall framework as it suits for binary classification tasks and under the scope of ISO standard. In the final implementation, FCN and FCN-OWT-UCM outperformed gPb-OWT-UCM and MRS in all the testing tiles.

6.2. Conclusions

The workflow of FCN-OWT-UCM proposed in this research is capable of extracting connected cadastral boundaries in complexed urban environment. Experiments carried out on UAV images of Muhoza achieved an average of 0.47 in precision (or 47% correctness), 0.52 in recall (or 52% completeness) and 0.49 in F-score. The results on the other study site, Busogo, achieved slightly lower result of around 0.44 in F-score. Very clean and clear boundaries were generated by the proposed method, avoiding the effect of messed building contours. In both study sites, the proposed method performed better than contending algorithms.

Relying on the strong feature learning and abstraction ability of FCN, we expand the field of automatic cadastral boundary mapping from rural to urban regions. The results of FCN are further processed by a grouping algorithm using OWT-UCM to obtain closed cadastral parcels. However, owing to the polarization of the feature map, the improvement is limited. We recommend to combine contour

detection techniques such as gPb to modify the result. So far, the technique is mainly suitable when a large proportion of boundaries are visible. The knowledge of the local experts is needed to correct the extracted boundaries and include them in a final cadastral system. We conclude that the proposed automated method followed by experts' final correction and verification can reduce processing time and labour force of the current cadastral mapping and data updating practices.

6.3. Recommendations

Based on current research, recommendations for future works are listed as follow:

1. This research focuses on automatic feature detection techniques for cadastral boundary mapping. Post-processing procedure to integrate the technique in real practices is not discussed. Therefore, designing a contextualized workflow for post-processing to better embed the automation result into cadastral mapping procedure is recommended.
2. Boundary classification is a binary classification task with highly imbalanced classes. The amount of boundary pixels is much less than non-boundaries. Therefore, giving boundary class more weight by using weighted loss function in FCNs is recommended to improve the classification accuracy.
3. Combining FCNs with state-of-the-art contour detection techniques such as gPb in cadastral boundary detection is also recommended. The output of gPb could serve as a low-level feature to feed FCN, which may generate a reasonable abstraction from the detected contours.

LIST OF REFERENCES

Ali, Z., & Ahmed, S. (2013). Extracting parcel boundaries from satellite imagery for a Land Information System. In *2013 6th International Conference on Recent Advances in Space Technologies (RAST)* (pp. 79–81). IEEE. https://doi.org/10.1109/RAST.2013.6581319

Arbeláez, P., Maire, M., Fowlkes, C., & Malik, J. (2011). Contour Detection and Hierarchical Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(5), 898–916. https://doi.org/10.1109/TPAMI.2010.161

Baatz, M., & Schäpe, A. (2000). *Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation*. Retrieved from http://www.ecognition.com/sites/default/files/405_baatz_fp_12.pdf

Babawuro, U., & Beiji, Z. (2012). Satellite Imagery Cadastral Features Extractions using Image Processing Algorithms : A Viable Option for Cadastral Science. *International Journal of Computer Science*, *9*(4), 30–38.

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(12), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615

Bergado, J. R., Persello, C., & Gevaert, C. (2016). A deep learning approach to the classification of sub-decimetre resolution aerial images. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1516–1519). IEEE. https://doi.org/10.1109/IGARSS.2016.7729387

Bowyer, K., Kranenburg, C., & Dougherty, S. (2001). Edge Detector Evaluation Using Empirical ROC Curves. *Computer Vision and Image Understanding*, *84*(1), 77–103. https://doi.org/10.1006/CVIU.2001.0931

Crommelinck, S., Bennett, R., Gerke, M., Nex, F., Yang, M. Y., & Vosselman, G. (2016). Review of automatic feature extraction from high-resolution optical sensor data for UAV-based cadastral mapping. *Remote Sensing*, *8*(8). https://doi.org/10.3390/rs8080689

Cunningham, K., Walker, G., Stahlke, E., Wilson, R., & Opportunity, U. M. (2011). Cadastral Audit and Assessments Using Unmanned Aerial Systems. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XXXVIII-1*/(C22), 213–216. https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-213-2011

Drăguţ, L., Csillik, O., Eisank, C., & Tiede, D. (2014). Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing*, *88*, 119–127. https://doi.org/10.1016/j.isprsjprs.2013.11.018

Drăguţ, L., Tiede, D., & Levick, S. R. (2010). ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, *24*(6), 859–871. https://doi.org/10.1080/13658810903174803

Enemark, S., Bell, C. K., Lemmen, C., & Mclaren, R. (2014). *Fit-For-Purpose land administration*. Copenhagen, Denmark. Retrieved from www.fig.net

Enemark, S., Mclaren, R., Lemmen, C., Antonio, D., Gitau, J., De Zeeuw, K., … Freccia, S. (2016). *Fit-For-Purpoes land administration: guiding principles for country implementation*. Retrieved from http://www.landcoalition.org/sites/default/files/documents/resources/gltn-landadministration.pdf

García-Pedrero, A., Gonzalo-Martín, C., & Lillo-Saavedra, M. (2017). A machine learning approach for agricultural parcel delineation through agglomerative segmentation. *International Journal of Remote Sensing*, *38*(7), 1809–1819. https://doi.org/10.1080/01431161.2016.1278312

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. Retrieved from http://arxiv.org/abs/1502.01852

Hossin, M., & Sulaiman, M. N. (2015). Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, *5*(2), 1–11. https://doi.org/10.5121/ijdkp.2015.5201

IAAO. (2015). *Standard on Digital Cadastral Maps and Parcel Identifiers*. Kansas City. Retrieved from https://www.iaao.org/media/standards/Standard_Digital_Cadastral_Maps_2015.pdf

Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. https://doi.org/10.1007/s13398-014-0173-7.2

ISO 19157. (2013). *ISO 19157:2013 - Geographic information -- Data quality*. Retrieved from https://www.iso.org/standard/32575.html

Koeva, M., Muneza, M., Gevaert, C., Gerke, M., & Nex, F. (2018). Using UAVs for map creation and updating. A case study in Rwanda. *Survey Review*, *50*(361), 312–325. https://doi.org/10.1080/00396265.2016.1268756

Li, Y., Wang, S., Tian, Q., & Ding, X. (2015). A survey of recent advances in visual feature detection. *Neurocomputing*, *149*(PB), 736–751. https://doi.org/10.1016/j.neucom.2014.08.003

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. Retrieved from https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Long_Fully_Convolutional_Networks_2015_CVPR_paper.html

Luo, X., Bennett, R. M., Koeva, M., & Lemmen, C. (2017). Investigating Semi-Automated Cadastral Boundaries Extraction from Airborne Laser Scanned Data. *Land*, *6*(3), 60. https://doi.org/10.3390/land6030060

Manyoky, M., Theiler, P., Steudler, D., & Eisenbeiss, H. (2012). Unmanned Aerial Vehicle in Cadastral Applications. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XXXVIII-1*/(September), 57–62. https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-57-2011

Martin, D. R., Fowlkes, C. C., & Malik, J. (2004). Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *26*(5), 530–549. https://doi.org/10.1109/TPAMI.2004.1273918

Maurice, M. J., Koeva, M. N., Gerke, M., Nex, F., & Gevaert, C. (2015). A photogrammetric approach for map updating using UAV in Rwanda. *GeoTechRwanda*, 1–8.

Musyoka, G. M. (2018). *Automatic Delineation of Small Holder Agricultural Field Boundaries Using Fully Convolutional Networks (MSc thesis)*. Enschede: University of Twente Faculty of Geo-Information and Earth Observation (ITC).

Nyandwi, E. (2018). *The Battle of Cadastral Intelligence: Measuring the Results of Competition between People and Machine in Creation of Cadastral Boundaries (MSc thesis)*. Enschede: University of Twente Faculty of Geo-Information and Earth Observation (ITC). Retrieved from https://webapps.itc.utwente.nl/librarywww/papers_2018/msc/la/nyandwi.pdf

Pal, N. R., & Pal, S. K. (1993). A review on image segmentation techniques. *Pattern Recognition*, *26*(9), 1277–1294. https://doi.org/10.1016/0031-3203(93)90135-J

Parida, P. K., Sanabada, M. K., & Tripathi, S. (2014). Cadastral resurvey using High Resolution Satellite ortho Image - Challenges: A case study in Odisha, India. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *40*(8), 1165–1170. https://doi.org/10.5194/isprsarchives-XL-8-1165-2014

Persello, C., & Stein, A. (2017). Deep Fully Convolutional Networks for the Detection of Informal Settlements in VHR Images. *IEEE Geoscience and Remote Sensing Letters*, *14*(12), 2325–2329. https://doi.org/10.1109/LGRS.2017.2763738

Ramadhani, S. A., Bennett, R. M., & Nex, F. C. (2018). Exploring UAV in Indonesian cadastral boundary data acquisition. *Earth Science Informatics*, *11*(1), 129–146. https://doi.org/10.1007/s12145-017-0314-6

Rubinov, E., Biraro, M., Fuller, S., & Bennett, R. (2015). From Barefoot to "Air-foot" surveyors. *GIM International*, *29*(6), 21–23.

Silva, M. A., & Stubkjær, E. (2002). A review of methodologies used in research on cadastral development. *Computers, Environment and Urban Systems*, *26*(5), 403–423. https://doi.org/10.1016/S0198-9715(02)00011-X

Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition, 1–14. https://doi.org/10.2146/ajhp170251

UNECE. (1996). *Land administration guidelines with special reference to countries in transition*. Geneva. Retrieved from http://www.unece.org/fileadmin/DAM/hlm/documents/Publications/land.administration.guidelines.e.pdf

Wassie, Y. A., Koeva, M. N., Bennett, R. M., & Lemmen, C. H. J. (2018). A procedure for semi-automated cadastral boundary feature extraction from high-resolution satellite imagery. *Journal of Spatial Science*, *63*(1), 75–92. https://doi.org/10.1080/14498596.2017.1345667

Williamson, I. (1997). The justification of cadastral systems in developing countries. *Geomatica*, *51*(1), 21–36. Retrieved from https://minerva-access.unimelb.edu.au/handle/11343/34003

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307

Appendix 1: The classification results of FCN on training tiles.

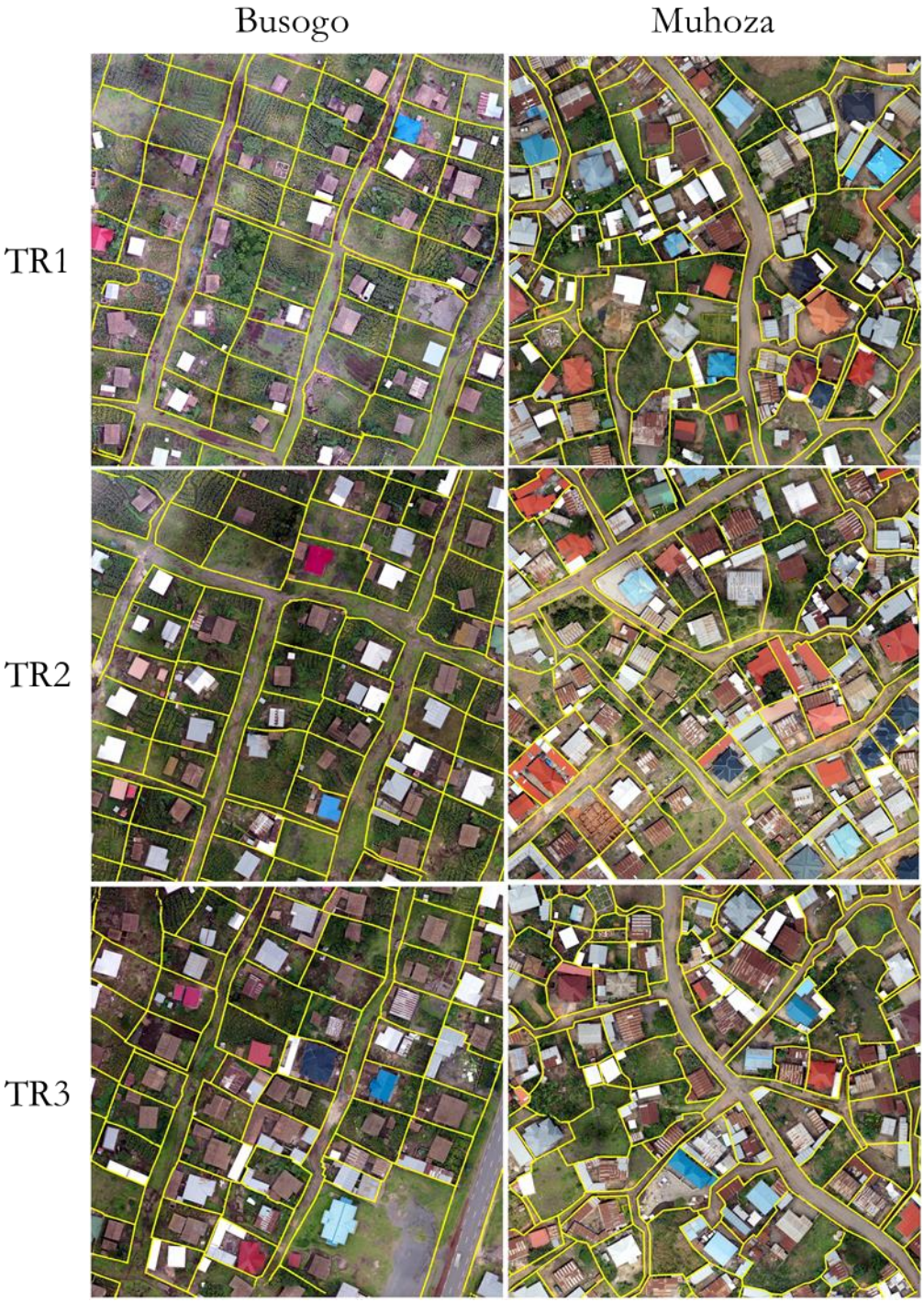


Figure 1: The extracted boundaries of FCN on training tiles in Busogo and Muhoza.

Table 1: The classification results of FCN on training tiles of Busogo and Muhoza.

Location	Tile	Precision	Recall	F-score
Busogo	TR1	0.9310	0.9386	0.9348
	TR2	0.9360	0.9405	0.9382
	TR3	0.9359	0.9414	0.9386
Muhoza	TR1	0.9499	0.9494	0.9497
	TR2	0.9568	0.9604	0.9586
	TR3	0.9515	0.9582	0.9548

Appendix 2: The threshold tuning records of gPb-OWT-UCM on testing tiles.

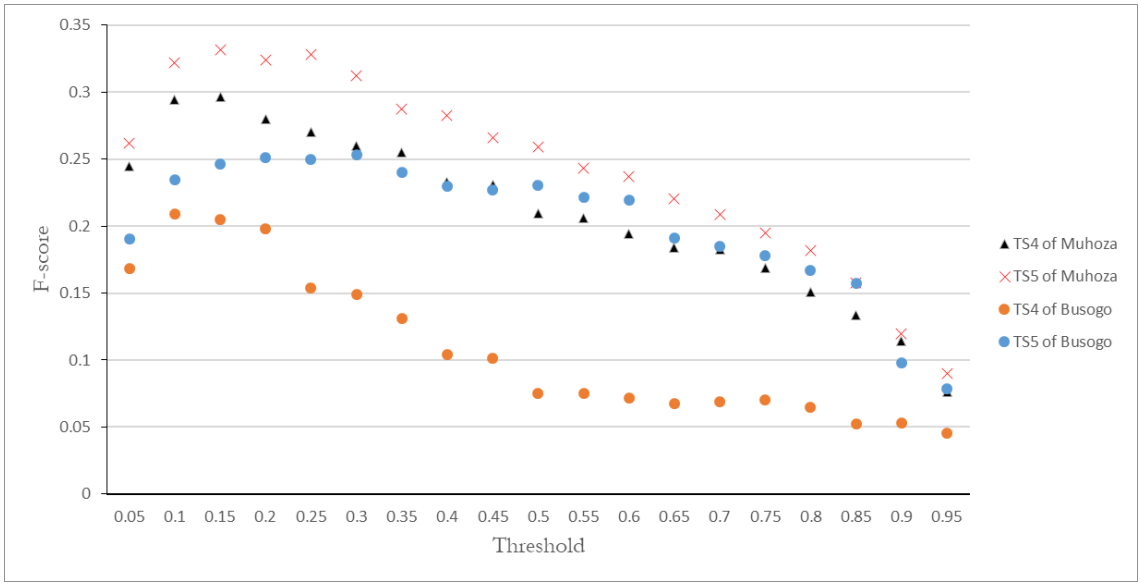


Figure 2: Threshold tuning results of gPb-OWT-UCM on the testing tiles.